J Glob Optim (2007) 37:357–379
DOI 10.1007/s10898-006-9053-9

ORIGINAL PAPER

Discrete global descent method for discrete global optimization and nonlinear integer programming

Chi-Kong Ng · Duan Li · Lian-Sheng Zhang

Received: 21 October 2005 / Acceptd: 2 June 2006 /
Published online: 17 August 2006

Abstract A novel method, entitled the discrete global descent method, is developed in this paper to solve discrete global optimization problems and nonlinear integer programming problems. This method moves from one discrete minimizer of the objective function f to another better one at each iteration with the help of an auxiliary function, entitled the discrete global descent function. The discrete global descent function guarantees that its discrete minimizers coincide with the better discrete minimizers of f under some standard assumptions. This property also ensures that a better discrete minimizer of f can be found by some classical local search methods. Numerical experiments on several test problems with up to 100 integer variables and up to 1.38×10^{104} feasible points have demonstrated the applicability and efficiency of the proposed method.

Keywords Discrete global descent method · Discrete global optimization · Nonlinear integer programming · Integer programming

1 Introduction

We consider in this paper the following discrete global optimization problem (nonlinear integer programming problem):

$$(P) \quad \min\{f(x) : x \in \mathbb{X}\}$$

with the following assumptions:

C.-K. Ng · D. Li (✉)
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong S.A.R., P.R.China.
e-mail: dli@se.cuhk.edu.hk

L.-S. Zhang
Department of Mathematics, Shanghai University, Baoshan, Shanghai 200436, P.R.China.

Assumption 1 $\mathbb{X} \subset \mathbb{Z}^n$ is a finite integer set with at least two integer points, where $\mathbb{Z}^n$ is the set of integer points in $\mathbb{R}^n$.

Assumption 1 implies that there exists a constant K such that

$$1 \leq \max_{x^{(1)},x^{(2)} \in \mathbb{X}} \|x^{(1)} - x^{(2)}\| \leq K < \infty, \tag{1}$$

where $\|\cdot\|$ is the usual Euclidean norm.

Assumption 2 $\mathbb{X}$ is a pathwise connected set (see Definition 1).

Assumption 3 $f : \mathbb{X} \mapsto \mathbb{R}$ satisfies the following Lipschitz condition for every $x^{(1)}$, $x^{(2)} \in \mathbb{X}$:

$$|f(x^{(1)}) - f(x^{(2)})| \leq L \, \|x^{(1)} - x^{(2)}\|, \tag{2}$$

where $0 < L < \infty$ is the Lipschitz constant.

Notice that the formulation in (P) allows the set $\mathbb{X}$ to be defined by box constraints as well as by inequality constraints. Furthermore, when f is coercive, i.e., $f(x) \to \infty$ as $\|x\| \to \infty$, there always exists a box which contains all discrete global minimizers of f. Thus, the unconstrained discrete global optimization problem, $\min\{f(x) : x \in \mathbb{Z}^n\}$ can be reduced into an equivalent problem formulation in (P). In other words, both unconstrained and constrained discrete global optimization problems can be considered in (P).

Discrete global optimization arises frequently in various applications such as combinatorics, computational finance, scheduling, design and operations problems. While a convexity in continuous optimization guarantees that a local search offers a global solution, this is certainly not the case for discrete optimization or integer programming. To support this argument, let us consider a two-dimensional example:

Example 1

$$\begin{array}{ll} \min & f(x) = (x - \bar{x})^T Q (x - \bar{x}) \\ \text{s.t.} & x \in \mathbb{X} = \{x \in \mathbb{Z}^2 : \; 0 \leq x_i \leq 7, \; i = 1, 2\}, \end{array}$$

where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \bar{x} = \begin{bmatrix} 3.1 \\ 2.5 \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} 42.67 & -49.41 \\ -49.41 & 57.38 \end{bmatrix}.$$

The global minimizer of this problem is $x^*_{\text{global}} = [6, 5]^T$ with $f(x^*_{\text{global}}) = 1$. Since matrix Q is positive definite, there is only one continuous local minimizer $\bar{x} = [3.1, 2.5]^T$ which is also the global minimizer. An integer point $x \in \mathbb{X}$ is defined here as a discrete local minimizer of f over $\mathbb{X}$ if its function value is less than or equal to that of its four neighboring points, $x \pm [1, 0]^T$ and $x \pm [0, 1]^T$, if they are included in $\mathbb{X}$. From Table 1 in the Appendix, it is clear that even for a convex function, there may exist multiple minima on an integer domain.

Wide applications of discrete global optimization and a discrepancy between integer local solutions and a global minimum solution to convex functions have made discrete global optimization an active and challenging research area in operations research and engineering.

During the last three decades, various deterministic solution methods have been proposed, e.g., branch and bound methods [11, 13], Lagrangian methods [5,9] nonlinear Lagrangian methods [14, 16, 25, 27], etc. However, these methods can only tackle relatively small scale general integer optimization problems.

The filled function methods are promising for continuous global optimization. The concept of the filled function together with the first filled function method were presented by Ge at the Tenth Biennial Conference on Numerical Analysis at Dundee, Scotland, in 1983, and his paper [6] was published in 1990. Thereafter, several filled functions with improved theoretical and computational properties better than the one in [6] have been proposed (see, e.g., [8, 17, 21, 26, 29]). Moreover, the filled function methods have been successfully applied in solving some practical problems (see, e.g., [1–3, 18]. Due to the promising results of the filled function methods for continuous global optimization, filled function approaches to discrete global optimization have been investigated since late 1980s. A survey of the utilization of the filled function methods in discrete global optimization can be found in [22].

The filled function approaches to discrete global optimization can be classified into two categories: continuous approaches and discrete approaches. Continuous approaches transform a discrete global optimization problem into a continuous global optimization problem and then solve it by some filled function methods. Methods using continuous approaches include [7] and [28]. As demonstrated in [22], the transformed objective function always generate more minima than the original objective function f. From the experience of continuous global optimization, a problem has more local minima is, in general, more difficult to be solved globally.

Discrete approaches search for a "discrete local minimizer" of f over $\mathbb{X}$ by a "gradient-free direct search method." Zhu [31] was probably the first one who used a filled function method in a discrete approach to discrete global optimization. He specifically used the traditional continuous filled function in [6] in his discrete approach. However, as demonstrated in [22], his method has some implementation problems. First, his discrete approach inherits the weakness of the continuous filled function in numerical implementation: the change in the filled function value is indistinguishable when the distance between the current iterative point, x, and the current local minimizer, x^*, being large. Moreover, a continuous filled function only guarantees that if f has a basin B_c^{**} that is lower than the current basin B_c^* at x^*, then there is a point x' in B_c^{**} that minimizes the continuous filled function on a straight line connecting x and x^*. These requirements are unlikely to be satisfied in discrete cases, not only because x' may not be an integer point, but also because not every point on the straight line is feasible in a discrete domain.

Recently, Ng et al. [23] proposed and formalized a discrete version of the filled function method, namely the discrete filled function method, for solving (P). Their method moves from one discrete local minimizer (see Definition 5) of f to another better one at each iteration with the help of an auxiliary function, namely the discrete filled function. The discrete filled function F_{x^*} ensures that the current discrete minimizer x^* of f is a strict discrete maximizer of F_{x^*}. Moreover, F_{x^*} has no discrete minimizer in the current discrete basin B^* at x^* or in any discrete basin of f higher than B^*. Furthermore, if f has a discrete basin B^{**} at x^{**} that is lower than B^*, then there is an integer point $x' \in B^{**}$ that minimizes F_{x^*} on a discrete path (see Definition 1) $\{x^*, \ldots, x', \ldots, x^{**}\}$ in $\mathbb{X}$. These properties are promising for finding such a transitional discrete point x' and in turn finding a better discrete minimizer x^{**} by some classical local search method, e.g., the discrete steepest descent method (see Algorithm 1). As witnessed in the numerical experiments reported in [23], the discrete filled function was efficient. Nevertheless, the theory of the discrete filled function only ensures that x' is a one-dimensional discrete minimizer over a discrete path. It does not guarantee x' is a true discrete minimizer over $\mathbb{X}$. This property introduces some numerical difficulties in the realization of the discrete filled function method in solving some practical problems.

In this paper, we enhance the concept of the discrete filled function and propose a family of sophisticated auxiliary functions, entitled the discrete global descent function, that not only keeps all the properties of a discrete filled function, but also ensures that x' coincides with x^{**} under some standard assumptions. This additional property guarantees that a better discrete minimizer of f, if it exists, can be found by some classical local search methods at each iteration.

This paper is organized as follows. Following this introduction, we present some preliminaries in Sect. 2 to streamline the discussion in this paper. In Sect. 3, we first give a formal definition of a discrete global descent function. We then propose a family of two-parameter discrete global descent functions and investigate its properties as well. After that, we consider the numerical implementation of the proposed discrete global descent method and suggest a solution algorithm in Sect. 4. In Sect. 5, we first demonstrate the solution procedures of the algorithm by an illustrative example. We then report the results of the algorithm in solving several test problems with up to 100 variables and up to 1.38×10^{104} feasible points. Finally, we draw some conclusions in Sect. 6.

2 Preliminaries

The aim of this section is to streamline the discussion in this paper. We recall some definitions and preliminary results in discrete analysis and discrete optimization.

2.1 Sets

Definition 1 A sequence $\{x^{(i)}\}_{i=0}^{u+1}$ is called a DISCRETE PATH in $\mathbb{X}$ between two distinct points x^* and x^{**} in $\mathbb{X}$ if $x^{(0)} = x^*, x^{(u+1)} = x^{**}, x^{(i)} \in \mathbb{X}$ for all i, $x^{(i)} \neq x^{(j)}$ for $i \neq j$ and $\|x^{(i+1)} - x^{(i)}\| = 1$ for all i. If such a discrete path exists, then x^* and x^{**} are said to be PATHWISE CONNECTED in $\mathbb{X}$. Furthermore, if every two distinct points in $\mathbb{X}$ are pathwise connected in $\mathbb{X}$, then $\mathbb{X}$ is called a PATHWISE CONNECTED SET.

Definition 2 The set of all axial directions in $\mathbb{Z}^n$ is defined by $\mathbb{D} = \{\pm e_i : i = 1, 2 \ldots, n\}$, where e_i is the ith unit vector (the n-dimensional vector with the ith component equal to one and all other components equal to zero).

Definition 3 For any $x \in \mathbb{Z}^n$, the DISCRETE NEIGHBORHOOD of x is defined by $N(x) = \{x,\ x \pm e_i : i = 1, 2, \ldots, n\}$.

Definition 4 A point $x \in \mathbb{X}$ is called a CORNER POINT of $\mathbb{X}$ if for each $d \in \mathbb{D}$, $x + d \in \mathbb{X}$ implies $x - d \notin \mathbb{X}$.

2.2 Discrete Optimization

Definition 5 A point $x^* \in \mathbb{X}$ is called a (DISCRETE) LOCAL MINIMIZER of f over $\mathbb{X}$ if $f(x^*) \leq f(x)$ for all $x \in \mathbb{X} \cap N(x^*)$. Furthermore, if $f(x^*) \leq f(x)$ for all $x \in \mathbb{X}$, then x^*

is called a (DISCRETE) GLOBAL MINIMIZER of f over $\mathbb{X}$. If, in addition, $f(x^*) < f(x)$ for all $x \in \mathbb{X} \cap N(x^*)\backslash x^*$, then x^* is called a STRICT (DISCRETE) LOCAL/GLOBAL MINIMIZER of f over $\mathbb{X}$.

Definition 6 For any $x \in \mathbb{X}$, $d \in \mathbb{D}$ is said to be a DESCENT DIRECTION of f at x over $\mathbb{X}$ if $x + d \in \mathbb{X}$ and $f(x + d) < f(x)$. Furthermore, $d^* \in \mathbb{D}$ is called a DISCRETE STEEPEST DESCENT DIRECTION of f at x over $\mathbb{X}$ if $f(x + d^*) \leq f(x + d)$ for any other descent direction d.

Similar to the continuous situation, we can design a discrete version of a steepest descent method for finding a local minimizer of f over $\mathbb{X}$.

Algorithm 1 *(Discrete Steepest Descent Method)*

0. Choose an initial point $x \in \mathbb{X}$.
1. If x is a local minimizer of f over $\mathbb{X}$, then stop. Otherwise, a discrete steepest descent direction $d^* \in \mathbb{D}$ of f at x over $\mathbb{X}$ can be found.
2. Let $x := x+\lambda d^*$, where $\lambda \in \mathbb{Z}_+$ is the step size such that f has a maximum decrease in the direction d^*. Go to Step 1.

2.3 Preliminary Results

The following basic properties will be useful in the later analysis. The proofs are not difficult and are thus omitted. Interested readers may find the proofs in [15, 21].

Lemma 1

(a) *For any $x^*, x^{**} \in \mathbb{X}$ and $d \in \mathbb{D}$, it holds $\|x^* - x^{**}\| \neq \|x^* + d - x^{**}\|$.*
(b) *For any $x^*, x^{**} \in \mathbb{X}$, if there exists $i \in \{1, 2, \ldots, n\}$ such that both $x^* \pm e_i \in \mathbb{X}$, then there exists $d \in \{\pm e_i\}$ such that $\|x^* + d - x^{**}\| > \|x^* - x^{**}\|$.*
(c) *If x^* and x^{**} are distinct strict local minimizers of f over $\mathbb{X}$, then $\|x^* - x^{**}\| > 1$.*

3 A family of discrete global descent functions and its properties

Denote by $\mathbb{X}^c$ the set of corner points of $\mathbb{X}$. Let x^* be a local minimizer of f over $\mathbb{X}$. Also, let $\hat{X}(x^*) = \{x \in \mathbb{X} : x \neq x^*, f(x) \geq f(x^*)\}$. We now give the formal definition for discrete global descent functions.

Definition 7 A function $G_{x^*} : \mathbb{X} \mapsto I\!R$ is said to be a DISCRETE GLOBAL DESCENT FUNCTION of f at x^* if it satisfies the following conditions:

(D1) x^* is a strict local maximizer of G_{x^*} over $\mathbb{X}$;
(D2) G_{x^*} has no local minimizer in the set $\hat{X}(x^*)\backslash\mathbb{X}^c$;
(D3) $x^{**} \in \mathbb{X}\backslash\mathbb{X}^c$ is a local minimizer of f over $\mathbb{X}$ with $f(x^{**}) < f(x^*)$ if and only if x^{**} is a local minimizer of G_{x^*} over $\mathbb{X}$.

We now propose a family of two-parameter discrete global descent functions for problem (P) at a local minimizer x^* of f over $\mathbb{X}$. Define

$$G_{\mu,\rho,x^*}(x) = A_\mu(f(x) - f(x^*)) - \rho\|x - x^*\|, \tag{3}$$

where $\rho > 0$ and $0 < \mu < 1$ are problem-dependent parameters,

$$A_\mu(y) = y \cdot V_\mu(y), \tag{4}$$

and $V_\mu : I\!R \mapsto I\!R$ is a continuous function that satisfies the following conditions:

(V1) $V_\mu(y)$ is strictly decreasing when $y < 0$ and non-increasing when $y \geq 0$,
(V2) $V_\mu(-\tau) = 1$, $V_\mu(0) = \mu$, and $V_\mu(y) \geq c\mu$ for all y,

where $\tau > 0$ is a sufficiently small number and $0 < c \leq 1$ is a constant. In theory, the parameter τ is required to satisfy:

$$0 < \tau < \min\{|f(x^{(1)}) - f(x^{(2)})| : x^{(1)}, x^{(2)} \in \mathbb{X},\ f(x^{(1)}) \neq f(x^{(2)})\}. \tag{5}$$

Thus, for any $x^{(1)}, x^{(2)} \in \mathbb{X}$, $f(x^{(1)}) < f(x^{(2)})$ implies $f(x^{(1)}) < f(x^{(2)}) - \tau$. If f is an integer-valued function over $\mathbb{X}$, we can simply set τ as a positive number less than 1. However, the discrete global descent algorithm developed in Sect. 4 is insensitive to the value of τ in numerical implementation. Therefore, τ is always set at 1 in calculation.

Some examples of V_μ that satisfy the above conditions are as follows.

Example 2 *Define*

$$V_\mu(y) = \begin{cases} (1-\mu)\left(\dfrac{y}{-\tau}\right)^{k+1} + \mu, & \text{if } y < 0 \\ \mu, & \text{if } y \geq 0 \end{cases}$$

for $k = 0, 1, 2, \ldots$.

Then $V_\mu \in C^k$ and V_μ satisfies conditions (V1) and (V2) with $c = 1$.

Example 3 Define

$$V^1_\mu(y) = \mu\left[(1-c)\left(\frac{1-c\mu}{\mu - c\mu}\right)^{-y/\tau} + c\right],$$

$$V^2_\mu(y) = \mu\left[\sqrt{(c'y)^2 + (1-c)^2} + c - c'y\right],$$

where $0 < c < 1$ and $c' = \dfrac{(1-\mu)(1+\mu-2c\mu)}{2\mu\tau(1-c\mu)}$.

It can be verified that both V^1_μ and $V^2_\mu \in C^\infty$ and satisfy conditions (V1) and (V2). Figure 1 in the Appendix illustrates $V^1_\mu(y)$, $V^2_\mu(y)$, $A^1_\mu(y) = y \cdot V^1_\mu(y)$ and $A^2_\mu(y) = y \cdot V^2_\mu(y)$ with $c = \mu = 0.5$ and $\tau = 1$.

The following lemma reveals some important properties of A_μ. These properties will be useful in the later analysis. The proofs are not difficult and are thus omitted. Interested readers may find the proofs in [15, 21].

Lemma 2

(a) $\operatorname{sgn}(A_\mu(y)) = \operatorname{sgn}(y)$.
(b) $A_\mu(y)$ *is a strictly increasing function of* y *for* $y \leq 0$.
(c) *If* $y^{(1)} < y^{(2)} \leq -\tau$, *then* $0 < y^{(2)} - y^{(1)} < A_\mu(y^{(2)}) - A_\mu(y^{(1)})$.
(d) *If* $y^{(1)} < -\tau < y^{(2)} < 0$, *then*
$A_\mu(y^{(1)}) < y^{(1)} < -\tau < y^{(2)} < A_\mu(y^{(2)}) < 0$.

In the next three Subsections (3.1–3.3), we will show that G_{μ,ρ,x^*} satisfies conditions (D1)–(D3) if the parameters μ and ρ satisfy certain conditions. We then provide an illustrative example for the proposed family of discrete global descent function in Subsection 3.4.

3.1 The proposed family of discrete global descent functions satisfies condition (D1)

Lemma 3 *Let* $\bar{x} \in \hat{X}(x^*)$. *If* $\rho > 0$ *and* $0 < \mu < \min\left\{1, \frac{\rho}{L}\right\}$, *then* $G_{\mu,\rho,x^*}(\bar{x}) < 0 = G_{\mu,\rho,x^*}(x^*)$.

Proof Since $f(\bar{x}) \geq f(x^*)$, by Assumption 3, we have $0 \leq f(\bar{x}) - f(x^*) \leq L\|\bar{x} - x^*\|$. Moreover, from the definition of V_μ, we have $V_\mu(y) \leq \mu$ for all $y \geq 0$. Thus, $V_\mu(f(\bar{x}) - f(x^*)) \leq \mu$. Therefore,

$$A_\mu(f(\bar{x}) - f(x^*)) = [f(\bar{x}) - f(x^*)] \cdot V_\mu(f(\bar{x}) - f(x^*)) \leq L\|\bar{x} - x^*\| \cdot \mu.$$

Since $\|\bar{x} - x^*\| > 0$, so if $\rho > 0$ and $0 < \mu < \min\left\{1, \frac{\rho}{L}\right\}$, we have

$$\begin{aligned} G_{\mu,\rho,x^*}(\bar{x}) &= A_\mu(f(\bar{x}) - f(x^*)) - \rho\|\bar{x} - x^*\| \\ &\leq L\mu\|\bar{x} - x^*\| - \rho\|\bar{x} - x^*\| < 0 = G_{\mu,\rho,x^*}(x^*). \end{aligned}$$

□

Theorem 4 *If* $\rho > 0$ *and* $0 < \mu < \min\left\{1, \frac{\rho}{L}\right\}$, *then* x^* *is a strict local maximizer of* G_{μ,ρ,x^*} *over* $\mathbb{X}$. *If, in addition,* x^* *is a global minimizer of* f *over* $\mathbb{X}$, *then* $G_{\mu,\rho,x^*}(x) < 0$ *for all* $x \in \mathbb{X}\backslash x^*$.

Proof Since x^* is a local minimizer of f over $\mathbb{X}$, thus $f(x) \geq f(x^*)$ for all $x \in \mathbb{X} \cap N(x^*)$. By Lemma 3, if $\rho > 0$ and $0 < \mu < \min\left\{1, \frac{\rho}{L}\right\}$, then $G_{\mu,\rho,x^*}(x) < 0 = G_{\mu,\rho,x^*}(x^*)$ for all $x \in \mathbb{X} \cap N(x^*)\backslash x^*$. Therefore, x^* is a strict local maximizer of G_{μ,ρ,x^*}.

If x^* is a global minimizer of f over $\mathbb{X}$, then $f(x) \geq f(x^*)$ for all $x \in \mathbb{X}$. The result then follows from Lemma 3. □

From Theorem 4, we conclude that G_{μ,ρ,x^*} satisfies the condition (D1) if $\rho > 0$ and $0 < \mu < \min\left\{1, \frac{\rho}{L}\right\}$.

3.2 The proposed family of discrete global descent functions satisfies condition (D2)

Lemma 5 *Let* $x^{(1)}$ *and* $x^{(2)}$ *be two integer points in* $\hat{X}(x^*)$ *such that* $0 < \|x^{(1)} - x^*\| < \|x^{(2)} - x^*\|$. *If* $\rho > 0$ *and* $0 < \mu < \min\left\{1, \frac{\rho}{(2K^2L)}\right\}$, *then*

$$G_{\mu,\rho,x^*}(x^{(2)}) < G_{\mu,\rho,x^*}(x^{(1)}) < 0 = G_{\mu,\rho,x^*}(x^*). \tag{6}$$

Proof We first show that

$$1 - \frac{\|x^{(1)} - x^*\|}{\|x^{(2)} - x^*\|} > \frac{1}{2K^2}. \tag{7}$$

Since $x^{(1)}, x^{(2)}$ and x^* are integer points and $\|x^{(1)} - x^*\| < \|x^{(2)} - x^*\|$, it holds

$$\|x^{(2)} - x^*\|^2 - \|x^{(1)} - x^*\|^2 \geq 1. \tag{8}$$

Moreover, by Assumption 1, we have $0 < \|x^{(2)} - x^*\| + \|x^{(1)} - x^*\| < 2K$. It then follows from (8) that

$$\|x^{(2)} - x^*\| - \|x^{(1)} - x^*\| \geq \frac{1}{\|x^{(2)} - x^*\| + \|x^{(1)} - x^*\|} > \frac{1}{2K}.$$

Dividing both side of the above inequality by $\|x^{(2)} - x^*\|$ and using $\|x^{(2)} - x^*\| \le K$ give rise to (7).

Since $f(x^{(2)}) \ge f(x^*)$, by Assumption 3, we have $0 \le f(x^{(2)}) - f(x^*) \le L\|x^{(2)} - x^*\|$. Moreover, from the definition of V_μ, we have $V_\mu(y) \le \mu$ for all $y \ge 0$. Thus, $V_\mu(f(x^{(2)}) - f(x^*)) \le \mu$ and

$$\begin{aligned} A_\mu(f(x^{(2)}) - f(x^*)) &= [f(x^{(2)}) - f(x^*)] \cdot V_\mu(f(x^{(2)}) - f(x^*)) \\ &\le L\|x^{(2)} - x^*\| \cdot \mu. \end{aligned}$$

Besides, since $f(x^{(1)}) \ge f(x^*)$, by Lemma 2(a), we have $A_\mu(f(x^{(1)}) - f(x^*)) \ge 0$. Therefore, by (7), if $\rho > 0$ and $0 < \mu < \min\left\{1, \frac{\rho}{(2K^2L)}\right\}$, then

$$\begin{aligned} G_{\mu,\rho,x^*}(x^{(2)}) - G_{\mu,\rho,x^*}(x^{(1)}) &= [A_\mu(f(x^{(2)}) - f(x^*)) - A_\mu(f(x^{(1)}) - f(x^*))] \\ &\quad - \rho(\|x^{(2)} - x^*\| - \|x^{(1)} - x^*\|) \\ &\le L\mu\|x^{(2)} - x^*\| - \rho(\|x^{(2)} - x^*\| - \|x^{(1)} - x^*\|) \\ &= \|x^{(2)} - x^*\| \cdot \left[L\mu - \rho\left(1 - \frac{\|x^{(1)} - x^*\|}{\|x^{(2)} - x^*\|}\right)\right] \\ &< \|x^{(2)} - x^*\| \cdot \left(L\mu - \frac{\rho}{2K^2}\right) < 0. \end{aligned}$$

By Assumption 1, we have $K \ge 1$, thus $0 < \mu < \min\left\{1, \rho/(2K^2L)\right\} \le \min\{1, \rho/L\}$. The second inequality of (6) follows directly from Lemma 3. □

Theorem 6 *Let $\bar{d} \in \mathbb{D}$ be a feasible direction at an integer point $\bar{x} \in \hat{X}(x^*)$ such that $\|\bar{x}+\bar{d}-x^*\| > \|\bar{x}-x^*\|$. If $\rho > 0$ and $0 < \mu < \min\left\{1, \rho/(2K^2L)\right\}$, then $G_{\mu,\rho,x^*}(\bar{x}+\bar{d}) < G_{\mu,\rho,x^*}(\bar{x}) < 0 = G_{\mu,\rho,x^*}(x^*)$.*

Proof Consider the following two cases:
Case (i): $f(\bar{x}+\bar{d}) \ge f(x^*)$. Since both $\bar{x}$ and $\bar{x}+\bar{d} \in \hat{X}(x^*)$, $0 < \|\bar{x}-x^*\| < \|\bar{x}+\bar{d}-x^*\|$, $\rho > 0$ and $0 < \mu < \min\left\{1, \rho/(2K^2L)\right\}$, it follows from Lemma 5 that $G_{\mu,\rho,x^*}(\bar{x}+\bar{d}) < G_{\mu,\rho,x^*}(\bar{x}) < 0 = G_{\mu,\rho,x^*}(x^*)$.
Case (ii): $f(\bar{x}+\bar{d}) < f(x^*) \le f(\bar{x})$. From Lemma 2(a), we have $A_\mu(f(\bar{x}+\bar{d}) - f(x^*)) < 0 \le A_\mu(f(\bar{x}) - f(x^*))$. Therefore, for $\rho > 0$,

$$\begin{aligned} G_{\mu,\rho,x^*}(\bar{x}+\bar{d}) &= A_\mu(f(\bar{x}+\bar{d}) - f(x^*)) - \rho\|\bar{x}+\bar{d}-x^*\| \\ &< A_\mu(f(\bar{x}) - f(x^*)) - \rho\|\bar{x}-x^*\| = G_{\mu,\rho,x^*}(\bar{x}). \end{aligned}$$

Since $0 < \mu < \min\left\{1, \rho/(2K^2L)\right\} \le \min\{1, \rho/L\}$, by Lemma 3, we have $G_{\mu,\rho,x^*}(\bar{x}+\bar{d}) < G_{\mu,\rho,x^*}(\bar{x}) < 0 = G_{\mu,\rho,x^*}(x^*)$. □

Corollary 7 *If $\rho > 0$ and $0 < \mu < \min\left\{1, \rho/(2K^2L)\right\}$, then G_{μ,ρ,x^*} satisfies the condition (D2).*

Proof For any $\bar{x} \in \hat{\mathbb{X}}(x^*) \setminus \mathbb{X}^c$, since $\bar{x}$ is not a corner point of $\mathbb{X}$, there exists $i \in \{1, 2, \ldots, n\}$ such that $\bar{x} \pm e_i \in \mathbb{X}$. By Lemma 1(b), there exists $\bar{d} \in \{\pm e_i\}$ such that $\|\bar{x}+\bar{d}-x^*\| > \|\bar{x}-x^*\|$. By Theorem 6, $\bar{d}$ is a feasible descent direction of G_{μ,ρ,x^*} at $\bar{x}$. Therefore, $\bar{x}$ is not a local minimizer of G_{μ,ρ,x^*}. □

3.3 The proposed family of discrete global descent functions satisfies condition (D3)

Theorem 8 *Let x^{**} be a strict local minimizer of f over $\mathbb{X}$ with $f(x^{**}) < f(x^*)$. If $\rho > 0$ is sufficiently small and $0 < \mu < 1$, then x^{**} is a strict local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$.*

Proof From Lemma 1(a), we have $\|x^{**} + d - x^*\| \neq \|x^{**} - x^*\|$ for all $d \in \mathbb{D}$. For any feasible direction $\bar{d} \in \mathbb{D}$ at x^{**}, we will show that

$$G_{\mu,\rho,x^*}(x^{**}) < G_{\mu,\rho,x^*}(x^{**} + \bar{d}). \tag{9}$$

Consider the following two cases:
Case (i): $\|x^{**} + \bar{d} - x^*\| < \|x^{**} - x^*\|$. If $f(x^{**}) < f(x^{**} + \bar{d}) \leq f(x^*)$, it then follows from Lemma 2(b) that

$$A_\mu(f(x^{**}) - f(x^*)) < A_\mu(f(x^{**} + \bar{d}) - f(x^*)). \tag{10}$$

Otherwise, if $f(x^{**}) < f(x^*) < f(x^{**} + \bar{d})$, from Lemma 2(a), we have

$$A_\mu(f(x^{**}) - f(x^*)) < 0 < A_\mu(f(x^{**} + \bar{d}) - f(x^*)). \tag{11}$$

Inequalities (10) and (11) imply that

$$\begin{aligned} G_{\mu,\rho,x^*}(x^{**}) &= A_\mu(f(x^{**}) - f(x^*)) - \rho\|x^{**} - x^*\| \\ &< A_\mu(f(x^{**} + \bar{d}) - f(x^*)) - \rho\|x^{**} + \bar{d} - x^*\| \\ &= G_{\mu,\rho,x^*}(x^{**} + \bar{d}). \end{aligned}$$

Case (ii): $\|x^{**} + \bar{d} - x^*\| > \|x^{**} - x^*\|$. By (5), $f(x^{**}) < f(x^*)$ implies $f(x^{**}) < f(x^*) - \tau$. Consider the following three cases:

$$f(x^{**}) < f(x^{**} + \bar{d}) \leq f(x^*) - \tau, \tag{12}$$

$$f(x^{**}) < f(x^*) - \tau < f(x^{**} + \bar{d}) < f(x^*), \tag{13}$$

$$f(x^{**}) < f(x^*) - \tau < f(x^*) \leq f(x^{**} + \bar{d}). \tag{14}$$

If (12) holds, from Lemma 2(c), we have

$$f(x^{**} + \bar{d}) - f(x^{**}) < A_\mu(f(x^{**} + \bar{d}) - f(x^*)) - A_\mu(f(x^{**}) - f(x^*)). \tag{15}$$

Let

$$\rho_1 = \min_{d \in \mathbb{D}_0(x^{**})} \frac{f(x^{**} + d) - f(x^{**})}{K}, \tag{16}$$

where $\mathbb{D}_0(x^{**}) = \{d \in \mathbb{D} : x^{**} + d \in \mathbb{X}\}$. Since x^{**} is a strict local minimizer of f over $\mathbb{X}$, we have $\rho_1 > 0$. Also, by Assumption 1 and Lemma 1(c), $\|x^{**} + \bar{d} - x^*\| - \|x^{**} - x^*\| < K$. Therefore, if $0 < \rho \leq \rho_1$, we obtain from (15) that

$$\begin{aligned} \rho \leq \rho_1 &\leq \frac{f(x^{**} + \bar{d}) - f(x^{**})}{K} \\ &< \frac{A_\mu(f(x^{**} + \bar{d}) - f(x^*)) - A_\mu(f(x^{**}) - f(x^*))}{\|x^{**} + \bar{d} - x^*\| - \|x^{**} - x^*\|}, \end{aligned}$$

which in turn implies (9).

If (13) holds, by Lemma 2(d), we have $A_\mu(f(x^{**})-f(x^*)) < f(x^{**})-f(x^*) < -\tau < f(x^{**}+\bar{d})-f(x^*) < A_\mu(f(x^{**}+\bar{d})-f(x^*)) < 0$. Therefore, (15) is satisfied and hence (9) holds if $0 < \rho < \rho_1$.

Finally, if (14) holds, since $f(x^{**}+\bar{d})-f(x^*) \geq 0$, by Lemma 2(a), we have

$$A_\mu(f(x^{**}+\bar{d})-f(x^*)) \geq 0. \tag{17}$$

Moreover, since $f(x^{**})-f(x^*) < -\tau$, by Lemma 2(d), we have

$$A_\mu(f(x^{**})-f(x^*)) < -\tau. \tag{18}$$

Let $\rho_2 = \frac{\tau}{K}$. If $0 < \rho \leq \rho_2$, by (17) and (18), we have

$$\begin{aligned}\rho \leq \frac{\tau}{K} &< \frac{\tau}{\|x^{**}+\bar{d}-x^*\| - \|x^{**}-x^*\|}\\ &< \frac{A_\mu(f(x^{**}+\bar{d})-f(x^*)) - A_\mu(f(x^{**})-f(x^*))}{\|x^{**}+\bar{d}-x^*\| - \|x^{**}-x^*\|}.\end{aligned}$$

Thus, (9) holds.

In summary, if $0 < \rho \leq \min\{\rho_1, \rho_2\}$, then x^{**} is a strict local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$. □

Theorem 8 assumes that the better local minimizer x^{**} of f over $\mathbb{X}$ is strict. This requirement on x^{**} can be relaxed to

$$f(x^{**}+\bar{d}) > f(x^{**}), \quad \text{for all } \bar{d} \in \mathbb{D}_1(x^{**}, x^*), \tag{19}$$

where $\mathbb{D}_1(x^{**}, x^*) = \{\, d \in \mathbb{D} : x^{**}+d \in \mathbb{X},\ \|x^{**}+d-x^*\| > \|x^{**}-x^*\| \,\}$.

Theorem 9 *Let x^{**} be a local minimizer of f over $\mathbb{X}$ with $f(x^{**}) < f(x^*)$ that satisfies* (19). *If $\rho > 0$ is sufficiently small and $0 < \mu < 1$, then x^{**} is a strict local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$.*

Proof From Lemma 1(a), we have $\|x^{**}+d-x^*\| \neq \|x^{**}-x^*\|$ for all $d \in \mathbb{D}$. Let $\bar{d} \in \mathbb{D}$ be a feasible direction at x^{**}. Since x^{**} is a local minimizer of f over $\mathbb{X}$, thus $f(x^{**}+\bar{d}) \geq f(x^{**})$. If, in addition, $\|x^{**}+\bar{d}-x^*\| > \|x^{**}-x^*\|$, by (19), we have $f(x^{**}+d^{(1)}) > f(x^{**})$. To prove (9), we can use the similar arguments as in the proof of Theorem 8 except for the following additional case: $\|x^{**}+\bar{d}-x^*\| < \|x^{**}-x^*\|$ and $f(x^{**}) = f(x^{**}+\bar{d}) < f(x^*)$. In this case, we have

$$\begin{aligned}G_{\mu,\rho,x^*}(x^{**}) &= A_\mu(f(x^{**})-f(x^*)) - \rho\|x^{**}-x^*\|\\ &< A_\mu(f(x^{**}+\bar{d})-f(x^*)) - \rho\|x^{**}+\bar{d}-x^*\|\\ &= G_{\mu,\rho,x^*}(x^{**}+\bar{d}).\end{aligned}$$

Thus, (9) holds. □

Theorem 9 clearly states that if there exists a pathwise connected set X^{**} such that every $x \in X^{**}$ is a local minimizer of f over $\mathbb{X}$ and $f(x) < f(x^*)$, then the farthest point $x^{**} \in X^{**}$ away from x^* is a strict local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$, provided that $\rho > 0$ is sufficiently small and $0 < \mu < 1$.

Theorem 10 *Let x' be a local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$ and $\bar{d} \in \mathbb{D}$ be a feasible direction at x' such that $\|x'+\bar{d}-x^*\| > \|x'-x^*\|$. If $\rho > 0$ is sufficiently small and $0 < \mu < \min\left\{1, \rho/(2K^2L)\right\}$, then x' is a local minimizer of f over $\mathbb{X}$.*

Proof Since x^* is a local minimizer of f over $\mathbb{X}$, by Theorem 4, x^* is a strict local maximizer of G_{μ,ρ,x^*}. Therefore, $x' \neq x^*$. We claim that $f(x') < f(x^*)$. Suppose on the contrary that $f(x') \geq f(x^*)$. Then $x' \in \hat{X}(x^*)$. Since $\|x' + \bar{d} - x^*\| > \|x' - x^*\|$, by Theorem 6, we have $G_{\mu,\rho,x^*}(x' + \bar{d}) < G_{\mu,\rho,x^*}(x')$, a contradiction to the assumption that x' is a local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$. Therefore, $f(x') < f(x^*)$ and $f(x') < f(x^*) - \tau$ by the definition of τ.

Now, suppose on the contrary that x' is not a local minimizer of f over $\mathbb{X}$. Then there exists a descent direction $d' \in \mathbb{D}$ at x' such that $f(x' + d') < f(x')$ and hence $f(x' + d') - f(x^*) < f(x') - f(x^*) < -\tau$. By Lemma 2(c), we have

$$0 < f(x') - f(x' + d') < A_\mu(f(x') - f(x^*)) - A_\mu(f(x' + d') - f(x^*)). \tag{20}$$

Since, from Lemma 1(a), $\|x' + d' - x^*\| \neq \|x' - x^*\|$. If $\|x' + d' - x^*\| > \|x' - x^*\|$, by (20), we have

$$\begin{aligned} G_{\mu,\rho,x^*}(x' + d') &= A_\mu(f(x' + d') - f(x^*)) - \rho\|x' + d' - x^*\| \\ &< A_\mu(f(x') - f(x^*)) - \rho\|x' - x^*\| = G_{\mu,\rho,x^*}(x'), \end{aligned}$$

which contradicts the assumption that x' is a local minimizer of G_{μ,ρ,x^*}.

On the other hand, if $\|x' + d' - x^*\| < \|x' - x^*\|$, we then choose ρ such that $0 < \rho \leq \rho_3$, where

$$\rho_3 = \min_{d \in \mathbb{D}_0(x')} \frac{f(x') - f(x' + d)}{K} > 0 \tag{21}$$

and $\mathbb{D}_0(x') = \{ d \in \mathbb{D} : x' + d \in \mathbb{X} \}$. By (20) and (21), we have

$$\begin{aligned} \rho &\leq \frac{f(x') - f(x' + d')}{K} \\ &< \frac{A_\mu(f(x') - f(x^*)) - A_\mu(f(x' + d') - f(x^*))}{\|x' - x^*\| - \|x' + d' - x^*\|}. \end{aligned}$$

This implies $G_{\mu,\rho,x^*}(x' + d') < G_{\mu,\rho,x^*}(x')$. Again, this is a contradiction. □

Corollary 11 *Assume that every local minimizer of f over $\mathbb{X}$ is strict. Suppose that $\rho > 0$ is sufficiently small and $0 < \mu < \min\{1, \rho/(2K^2L)\}$. Then, $x^{**} \in \mathbb{X}\backslash\mathbb{X}^c$ is a local minimizer of f over $\mathbb{X}$ with $f(x^{**}) < f(x^*)$ if and only if x^{**} is a local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$.*

Proof The "if" part follows directly from Theorem 8. Now, suppose that x^{**} is a local minimizer of G_{μ,ρ,x^*} over $\mathbb{X}$. Since $x^{**} \notin \mathbb{X}^c$, there exists $i \in \{1, 2, \ldots, n\}$ such that both $x^{**} \pm e_i \in \mathbb{X}$. Thus, by Lemma 1(b), there exists a feasible direction $d \in \{\pm e_i\}$ at x^{**} such that $\|x^{**} + d - x^*\| > \|x^{**} - x^*\|$. If $\rho > 0$ is small enough and $0 < \mu \leq \min\{1, \rho(2K^2L)\}$, by Theorem 10, x^{**} is a local minimizer of f over $\mathbb{X}$. □

Corollary 11 indicates that if every local minimizer of f over $\mathbb{X}$ is strict, then G_{μ,ρ,x^*} satisfies the condition (D3) for suitable parameters μ and ρ.

3.4 An illustration of the proposed family of discrete global descent functions

We consider now the following illustrative example.

Example 4 (3-hump back camel function)

$$\begin{aligned} \min\ & f(x) = 2x_1^2 - 1.05x_1^4 + \frac{1}{6}x_1^6 - x_1x_2 + x_2^2, \\ \text{s.t.}\ & x_i = y_i/1000, \quad i = 1, 2, \\ & -2000 \le y_1 \le 2000, \quad -1500 \le y_2 \le 1500, \quad y_1, y_2 \text{ integers.} \end{aligned}$$

This problem has three local minima: $x^{*1} = [-1.748, -0.874]^T$ with $f(x^{*1}) = 0.2986$, $x^{*2} = [1.748, 0.874]^T$ with $f(x^{*2}) = 0.2986$ and $x^{*3} = [0, 0]^T$ with $f(x^{*3}) = 0$, among which x^{*3} is the global optimal solution. Let $x^* = x^{*2}$ be the current local minimizer. We construct a discrete global descent function G_{μ,ρ,x^*} at x^* with $\mu = \rho = 0.01$. Figure 2 in the Appendix shows the contours of f and G_{μ,ρ,x^*} and the figures of f and G_{μ,ρ,x^*}.

4 Numerical implementation and solution algorithm

Based on the theoretical results in the previous section, the discrete global descent method for (P) is described now as follows.

Algorithm 2 *(Discrete Global Descent Method for DGO and NLIP)*

0. (Initialization).
   - (i) Choose a function V_μ satisfying conditions (V1) and (V2).
   - (ii) Choose an initial point $x_{\text{ini}} \in \mathbb{X}$, a lower bound of ρ: $\rho_L > 0$, and two fractions: $\hat{\rho}$ $(0 < \hat{\rho} < 1)$ and $\hat{\mu}$ $(0 < \hat{\mu} < 1)$.
   - (iii) Starting from x_{ini}, apply Algorithm 1 to obtain a local minimizer x^* of f over $\mathbb{X}$. Set $k := 0$.
1. Generate a set of m initial points: $\{ x_{\text{ini}}^{(i)} \in \mathbb{X} \backslash x^* : i = 1, 2, \ldots, m \}$. Set $i := 1$.
2. Set the current point $x_{\text{cur}} := x_{\text{ini}}^{(i)}$.
3. If $f(x_{\text{cur}}) < f(x^*)$, then starting from x_{cur}, apply Algorithm 1 to find a local minimizer x^{**} of f over $\mathbb{X}$ such that $f(x^{**}) < f(x^*)$. Set $x^* := x^{**}$, $k := k + 1$. Go to Step 1.
4. Let $\mathbb{D}_0 := \{ d \in \mathbb{D} : x_{\text{cur}} + d \in \mathbb{X} \}$. If there exists $d \in \mathbb{D}_0$ such that $f(x_{\text{cur}} + d) < f(x^*)$, then starting from $x_{\text{cur}} + d^*$, where $d^* = \arg\min\{ f(x_{\text{cur}} + d) : d \in \mathbb{D}_0 \}$, apply Algorithm 1 to find a local minimizer x^{**} such that $f(x^{**}) < f(x^*)$. Set $x^* := x^{**}$, $k := k + 1$. Go to Step 1.
5. If x_{cur} is a local minimizer of G_{μ,ρ,x^*} and the set $\mathbb{D}_1 := \{ d \in \mathbb{D}_0 : \|x_{\text{cur}} + d - x^*\| > \|x_{\text{cur}} - x^*\| \}$ is empty, then go to Step 8.
6. If x_{cur} is a local minimizer of G_{μ,ρ,x^*}, then set $\mu_0 := \mu$ and choose a positive integer l such that $\mu = \hat{\mu}^l \mu_0$ and there exists a descent direction of G_{μ,ρ,x^*} at x_{cur}.
7. Let $\mathbb{D}_2 := \{ d \in \mathbb{D}_0 : G_{\mu,\rho,x^*}(x_{\text{cur}} + d) < G_{\mu,\rho,x^*}(x_{\text{cur}}),\ f(x_{\text{cur}} + d) < f(x_{\text{cur}}) \}$. If $\mathbb{D}_2 \ne \emptyset$, then set
$$d^* := \arg\min\{ f(x_{\text{cur}} + d) + G_{\mu,\rho,x^*}(x_{\text{cur}} + d) : d \in \mathbb{D}_2 \}.$$
   Otherwise set
$$d^* := \arg\min\{ G_{\mu,\rho,x^*}(x_{\text{cur}} + d) : d \in \mathbb{D}_0 \}.$$
   Set $x_{\text{cur}} := x_{\text{cur}} + d^*$. Go to Step 4.
8. Set $i := i + 1$. If $i \le m$, go to Step 2.

9. Set $\rho := \hat{\rho}\rho$. If $\rho \geq \rho_L$, then go to Step 1. Otherwise, the algorithm is incapable of finding a better local minimizer starting from the initial points, $\{x_{\text{ini}}^{(i)} : i = 1, 2, \ldots, m\}$. The algorithm stops and x^* is taken as a global minimizer.

The motivation and mechanism behind the algorithm are explained below.

A set of m initial points is generated in Step 1 to minimize G_{μ,ρ,x^*}. If no additional information about the objective function is provided, we set the initial points symmetric about the current local minimizer. For example, we can set $m = 2n$ and choose $x^* \pm e_i$, for $i = 1, 2, \ldots, n$, as initial points for the discrete global descent method.

Step 3 represents the situation where the current initial point, $x_{\text{ini}}^{(i)}$, satisfies $f(x_{\text{ini}}^{(i)}) < f(x^*)$. Therefore, we can further minimize the objective function f by any discrete local minimization method starting from $x_{\text{ini}}^{(i)}$. Note that, Step 3 is necessary only if we choose some initial points outside the discrete neighborhood of x^*.

Recall from Theorem 6 that if $f(x_{\text{cur}}) \geq f(x^*)$ and μ is sufficiently small, then x_{cur} cannot be a local minimizer of G_{μ,ρ,x^*}. In determining whether the current point x_{cur} is a local minimizer of G_{μ,ρ,x^*}, we compare $G_{\mu,\rho,x^*}(x_{\text{cur}})$ with $G_{\mu,\rho,x^*}(x)$ for all $x \in \mathbb{X} \cap N(x_{\text{cur}}) \backslash x_{\text{cur}}$. Step 4 represents the situation where one of the feasible neighboring points of x_{cur}, namely $x_{\text{cur}} + d^*$ with $d^* \in \mathbb{D}$, has a smaller objective function value than the current local minimum. We can then further minimize f by any discrete local minimization method starting from $x_{\text{cur}} + d^*$.

If it is found that x_{cur} is a local minimizer of G_{μ,ρ,x^*} with $f(x_{\text{cur}}) \geq f(x^*)$, this implies that μ is not small enough. Step 5 represents the situation when it is impossible to move further away from x^* than x_{cur} and thus x_{cur} must be a corner point of $\mathbb{X}$. Then, we give up the point x_{cur} without reducing the value of μ and try another initial point generated in Step 1. On the other hand, if x_{cur} is not a corner point of $\mathbb{X}$, then Step 6 reduces the value of μ to a pre-selected fraction recursively until there exists a descent direction of G_{μ,ρ,x^*} at x_{cur}.

Step 7 aims at selecting a more promising successor point. Note that if the algorithm goes from Step 6 to Step 7, G_{μ,ρ,x^*} has at least one descent direction at x_{cur}. If there exists a descent direction of both f and G_{μ,ρ,x^*} at x_{cur}, we then reduce both f and G_{μ,ρ,x^*} at the same time in order to take advantages of their reductions. On the other hand, if every descent direction of G_{μ,ρ,x^*} at x_{cur} is an increasing direction of f at x_{cur}, we reduce G_{μ,ρ,x^*} alone.

Recall from Corollary 11 that the value of ρ should be selected small enough. Otherwise, there could not exist a local minimizer of G_{μ,ρ,x^*}, even there exists a better x^{**} with $f(x^{**}) < f(x^*)$. Thus, the value of ρ is reduced successively in the solution process in Step 9 if no better solution is found when minimizing the discrete global descent function. If the value of ρ reaches its lower bound ρ_L and no better solution is found, the current local minimizer is taken as a global minimizer.

5 Numerical experiment

The developed discrete global descent method is programmed in Matlab 6.5 Release 13 and run on a Pentium IV system with 3.2 GHz CPU. An illustrative example is given first in the following to show the solution procedure of the algorithm described in the previous section. The computational results in solving several test problems are then reported.

Through out the tests, $V^1_\mu(y)$ (in Example 3) with $\tau = 1$ and $c = 0.5$ is selected for the discrete global descent function. Algorithm 1 is used to perform local searches. Suppose that a local minimizer x^{**} of f over $\mathbb{X}$ is obtained using $x^* + e_j$ as the initial point, the neighboring points of x^{**} are then arranged in the following order as the initial points in minimizing the discrete global descent function:

$$\begin{gathered} x^{**} + e_j, \\ x^{**} + e_{j+1}, x^{**} - e_{j+1}, \ldots, x^{**} + e_n, x^{**} - e_n, \\ x^{**} + e_1, x^{**} - e_1, \ldots, x^{**} + e_{j-1}, x^{**} - e_{j-1}, \\ x^{**} - e_j. \end{gathered}$$

Notice that, if the current local minimizer of f is on the boundary of $\mathbb{X}$, then there are less than $2n$ initial points. In addition, $\rho = \rho_L = 0.1$ is set in all the tests. In other words, if the algorithm could not find a local minimizer of G_{μ,ρ,x^*} using all initial points, the algorithm stops immediately. Besides these, $\mu = 0.1$ is set at the beginning of the algorithm. Once the current μ is classified as insufficiently small, μ is reduced to $\mu/10$.

In each test problem, we will first give a mathematical model. After that, we will summarize the computational results of the discrete global descent method with the number of runs of the algorithm (N_{test}), the number of iteration cycles (N_{iter}), the CPU time in seconds to obtain the final results (T_{final}), the CPU time in seconds for the algorithm to stop at Step 9 (T_{stop}), the total number of objective function evaluations to obtain the final results (N_{final}), and the total number of objective function evaluations to stop at Step 9 (N_{stop}).

We consider now the following illustrative example.

Example 5 ([7, 28)]

$$\begin{aligned} \min \quad & f(x) = x_1 + 10x_2 \\ \text{s.t.} \quad & 66x_1 + 14x_2 \geq 1430, \\ & -82x_1 + 28x_2 \geq 1306, \\ & 0 \leq x_1 \leq 15, \quad 68 \leq x_2 \leq 102, \quad x_1, x_2 \text{ integers.} \end{aligned}$$

This problem has 314 feasible points. The global minimum solution is $x^*_{\text{global}} = [7, 70]^T$ with $f(x^*_{\text{global}}) = 707$.

The algorithm starts from a feasible point $x_{\text{ini}} = [15, 102]^T$ with $f(x_{\text{ini}}) = 1035$. By the discrete steepest descent method, an initial local minimizer $x^* = [3, 88]^T$ is obtained with $f(x^*) = 883$.

In the first iteration of the algorithm, $\mu = 0.1$ is found to be not small enough. When $\mu = 0.01$, the algorithm starts from $x^{(1)}_{\text{ini}} = [4, 88]^T$ and reaches $x' = [4, 87]^T$ with $f(x') = 874 < f(x^*)$. Then, the algorithm switches to the local search again and obtains $x^{**} = [4, 84]^T$ with $f(x^{**}) = 844$.

In the second iteration of the algorithm, the algorithm sets $x^* = [4, 84]^T$ and starts from $x^{(1)}_{\text{ini}} = [5, 84]^T$ and reaches $x' = [5, 83]^T$ with $f(x') = 835 < f(x^*)$. Then, the algorithm switches to the local search and obtains $x^{**} = [5, 79]^T$ with $f(x^{**}) = 795$.

In the same fashion, the algorithm generates $x^* = [5, 79]^T$, $x^{(1)}_{\text{ini}} = [6, 79]^T$, $x' = [6, 78]^T$ with $f(x') = 786 < f(x^*)$, $x^{**} = [6, 74]^T$ with $f(x^{**}) = 746$ in the third iteration. Similarly, the algorithm generates $x^* = [6, 74]^T$, $x^{(1)}_{\text{ini}} = [7, 74]^T$, $x' = [7, 73]^T$ with $f(x') = 737 < f(x^*)$, $x^{**} = [7, 70]^T$ with $f(x^{**}) = 707$ in the fourth iteration. The cumulative number of function evaluations is 79.

In the fifth iteration of the algorithm, three starting points, $[8,70]^T$, $[6,70]^T$ and $[7,69]^T$ are infeasible. Besides these, the algorithm cannot find a feasible point with function value less than 707 using the remaining starting point $[7,71]^T$. The cumulative number of function evaluations is 193.

In general, ρ should be reduced by a fraction and continue the process until $\rho < \rho_L$. Since $\rho = \rho_L = 0.1$ is selected in the numerical tests, and thus the algorithm is terminated. Therefore, $N_{\text{iter}} = 4$, $x^*_{\text{global}} = [7,70]^T$, $f(x^*_{\text{global}}) = 707$, $N_{\text{final}} = 79$ and $N_{\text{stop}} = 193$. The ratio of the number of function evaluations to reach the global minimum to the number of feasible points is $\frac{79}{314} \approx 0.2516$.

The following test problems are used in the computational experiments in testing the discrete global descent method.

Problem 1 ([4, 19])

$$\begin{aligned}
\min \;& f(x) = x_1^2 + x_2^2 + 3x_3^2 + 4x_4^2 + 2x_5^2 - 8x_1 - 2x_2 - 3x_3 - x_4 - 2x_5,\\
\text{s.t.} \;& x_1 + x_2 + x_3 + x_4 + x_5 \le 400,\\
& x_1 + 2x_2 + 2x_3 + x_4 + 6x_5 \le 800,\\
& 2x_1 + x_2 + 6x_3 \le 200,\\
& x_3 + x_4 + 5x_5 \le 200,\\
& x_1 + x_2 + x_3 + x_4 + x_5 \ge 55,\\
& x_1 + x_2 + x_3 + x_4 \ge 48,\\
& x_2 + x_4 + x_5 \ge 34,\\
& 6x_1 + 7x_5 \ge 104,\\
& 0 \le x_i \le 99, \quad x_i \text{ integer}, \quad i = 1,2,3,4,5.
\end{aligned}$$

This problem has 251401581 feasible points. The global minimum solution is $x^*_{\text{global}} = [16,22,5,5,7]^T$ with $f(x^*_{\text{global}}) = 807$. Five initial points were used in the test experiment: $x_{\text{ini}} = [17,18,7,7,9]^T$, $[21,34,0,0,0]^T$, $[0,0,0,48,15]^T$, $[100,0,0,0,40]^T$ and $[0,8,32,8,32]^T$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratio of the average number of function evaluations to reach the global minimum to the number of feasible points was about 1.41×10^{-5}. A summary of the computational results is displayed in Table 2 in the Appendix.

Problem 2 (*Colville's function* 4 [12, 24])

$$\begin{aligned}
\min \;& f(x) = 100\left(x_2 - x_1^2\right)^2 + (1-x_1)^2 + 90\left(x_4 - x_3^2\right)^2 + (1-x_3)^2\\
& \qquad + 10.1[(x_2-1)^2 + (x_4-1)^2] + 19.8(x_2-1)(x_4-1),\\
\text{s.t.} \;& -10 \le x_i \le 10, \quad x_i \text{ integer}, \quad i = 1,2,3,4.
\end{aligned}$$

The global minimum solution is $x^*_{\text{global}} = [1,1,1,1]^T$ with $f(x^*_{\text{global}}) = 0$. Nine initial points were used in the test experiment: $x_{\text{ini}} = [\alpha,\alpha,\alpha,\alpha]^T$ for $\alpha = 0, \pm 5, \pm 10$, and $x_{\text{ini}} = [\beta,\beta,-\beta,-\beta]^T$ and $[\beta,-\beta,\beta,-\beta]^T$ for $\beta = \pm 10$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratio of the average number of function evaluations to reach the global minimum to the number of feasible points was about 8.36×10^{-3}. A summary of the computational results is displayed in Table 3 in the Appendix.

Problem 3 (*Goldstein and Price's function* [10, 30])

$$\min\ f(x) = g(x)h(x),$$
$$\text{s.t.}\ \ x_i = y_i/1000, \quad -2000 \le y_i \le 2000, \quad y_i \text{ integer}, \quad i = 1,2,$$

where

$$g(x) = 1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2),$$
$$h(x) = 30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2).$$

The global minimum solution is $x^*_{\text{global}} = [0, -1]^T$ with $f(x^*_{\text{global}}) = 3$. Seven initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \alpha]^T$ for $\alpha = 0, \pm 1, \pm 2$, and $x_{\text{ini}} = [\beta, -\beta]^T$ for $\beta = \pm 2$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratio of the average number of function evaluations to reach the global minimum to the number of feasible points was about 5.32×10^{-4}. A summary of the computational results is displayed in Table 4 in the Appendix.

Problem 4 (*Beale's function* [20, 24])

$$\min\ f(x) = \left[1.5 - x_1(1 - x_2)\right]^2 + \left[2.25 - x_1\left(1 - x_2^2\right)\right]^2$$
$$+ \left[2.625 - x_1\left(1 - x_2^3\right)\right]^2,$$
$$\text{s.t.}\ \ x_i = y_i/1000, \quad -10^4 \le y_i \le 10^4, \quad y_i \text{ integer}, \quad i = 1,2.$$

The global minimum solution is $x^*_{\text{global}} = [3, 0.5]^T$ with $f(x^*_{\text{global}}) = 0$. Seven initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \alpha]^T$ for $\alpha = 0, \pm 5, \pm 10$, and $x_{\text{ini}} = [\beta, -\beta]^T$ for $\beta = \pm 10$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratio of the average number of function evaluations to reach the global minimum to the number of feasible points was about 4.08×10^{-4}. A summary of the computational results is displayed in Table 5 in the Appendix.

Problem 5 *(A combination of problems* 231 *and* 233 *in* [24])

$$\min\ f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$
$$\text{s.t.}\ \ x_1^2 + x_2^2 \ge 0.25, \quad -\frac{1}{3}x_1 + x_2 \ge -0.1,$$
$$x_i = y_i \times 10^{-4}, \quad 0 \le y_i \le 10^5, \quad y_i \text{ integer}, \quad i = 1,2.$$

This problem has about 8.413×10^9 feasible points. The global minimum solution is $x^*_{\text{global}} = [1, 1]^T$ with $f(x^*_{\text{global}}) = 0$. Nine initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \alpha]^T$ for $\alpha = 2, 4, 6, 8, 10$, and $x_{\text{ini}} = [0, 0.5]^T$, $[0, 10]^T$, $[10, 3.2334]^T$ and $[0.3536, 0.3536]^T$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratio of the average number of function evaluations to reach the global minimum to the number of feasible points was about 6.26×10^{-5}. A summary of the computational results is displayed in Table 6 in the Appendix.

Problem 6 (*Powell's singular function* [20, 24])

$$\begin{aligned} &\min\ f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4, \\ &\text{s.t.}\ \ x_i = y_i/1000, \quad -10^4 \le y_i \le 10^4, \quad y_i \text{ integer}, \quad i = 1, 2, 3, 4. \end{aligned}$$

The global minimum solution is $x^*_{\text{global}} = [0, \ldots, 0]^T$ with $f(x^*_{\text{global}}) = 0$. Ten initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \alpha, \alpha, \alpha]^T$ for $\alpha = \pm 3, \pm 6, \pm 10$, and $x_{\text{ini}} = [\beta, \beta, -\beta, -\beta]^T$ and $[\beta, -\beta, \beta, -\beta]^T$ for $\beta = \pm 10$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratio of the average number of function evaluations to reach the global minimum to the number of feasible points was about 2.78×10^{-11}. A summary of the computational results is displayed in Table 7 in the Appendix.

Problem 7 ([24])

$$\begin{aligned} &\min\ f(x) = (x_1 - 1)^2 + (x_n - 1)^2 + n \sum_{i=1}^{n-1} (n - i)(x_i^2 - x_{i+1})^2, \\ &\text{s.t.}\ \ -5 \le x_i \le 5, \quad x_i \text{ integer}, \quad i = 1, 2, \ldots, n. \end{aligned}$$

The global minimum solution is $x^*_{\text{global}} = [1, \ldots, 1]^T$ with $f(x^*_{\text{global}}) = 0$ for all n. Three sizes of the problem were considered: $n = 25, 50$ and 100. For all problems with different sizes, nine initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \ldots, \alpha]^T$ for $\alpha = 0, \pm 3, \pm 5$, and $x_{\text{ini}} = [\beta, \ldots, \beta, \beta, -\beta, \ldots, -\beta]^T$ and $[\beta, -\beta, \beta, -\beta, \ldots]^T$ for $\beta = \pm 5$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratios of the average numbers of function evaluations to reach the global minima to the numbers of feasible points were about 2.30×10^{-23}, 8.27×10^{-49} and 2.74×10^{-100}, for $n = 25, 50$ and 100, respectively. A summary of the computational results is displayed in Table 8 in the Appendix.

Problem 8 (*Rosenbrock's function* [24])

$$\begin{aligned} &\min\ f(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2], \\ &\text{s.t.}\ \ -5 \le x_i \le 5, \quad x_i \text{ integer}, \quad i = 1, 2, \ldots, n. \end{aligned}$$

The global minimum solution is $x^*_{\text{global}} = [1, \ldots, 1]^T$ with $f(x^*_{\text{global}}) = 0$ for all n. Three sizes of the problem were considered: $n = 25$, 50 and 100. For all problems with different sizes, nine initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \ldots, \alpha]^T$ for $\alpha = 0, \pm 3, \pm 5$, and $x_{\text{ini}} = [\beta, \ldots, \beta, \beta, -\beta, \ldots, -\beta]^T$ and $[\beta, -\beta, \beta, -\beta, \ldots]^T$ for $\beta = \pm 5$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratios of the average numbers of function evaluations to reach the global minima to the numbers of feasible points were about 8.31×10^{-22}, 6.21×10^{-47} and 4.27×10^{-98}, for $n = 25, 50$ and 100, respectively. A summary of the computational results is displayed in Table 9 in the Appendix.

Problem 9 ([23])

$$\begin{aligned} &\min\ f(x) = \sum_{i=1}^{n} x_i^4 + \left(\sum_{i=1}^{n} x_i \right)^2 \\ &\text{s.t.}\ \ -5 \le x_i \le 5, \quad x_i \text{ integer}, \quad i = 1, 2, \ldots, n. \end{aligned}$$

The global minimum solution is $x^*_{\text{global}} = [0, \ldots, 0]^T$ with $f(x^*_{\text{global}}) = 0$ for all n. Three sizes of the problem were considered: $n = 25, 50$ and 100. For all problems with different sizes, ten initial points were used in the test experiment: $x_{\text{ini}} = [\alpha, \ldots, \alpha]^T$ for $\alpha = \pm 1, \pm 3, \pm 5$, and $x_{\text{ini}} = [\beta, \ldots, \beta, \beta, -\beta, \ldots, -\beta]^T$ and $[\beta, -\beta, \beta, -\beta, \ldots]^T$ for $\beta = \pm 5$. For every experiment, the discrete global descent method succeeded in identifying the global minimum solution. The ratios of the average numbers of function evaluations to reach the global minima to the numbers of feasible points were about 8.90×10^{-22}, 2.00×10^{-47} and 6.60×10^{-99}, for $n = 25$, 50 and 100, respectively. A summary of the computational results is displayed in Table 10 in the Appendix.

6 Concluding remarks

A discrete global descent method for discrete global optimization and nonlinear integer programming has been developed in this paper. The sophisticated family of discrete global descent functions not only guarantees to have a local minimizer over the problem domain, but also ensures that its local minimizers coincide with the better local minimizers of f. This property assures that a local minimizer can be found by using some classical local search methods. Promising computation results have been observed from our numerical experiments of large scale. These results indicate the efficiency of the proposed method.

Acknowledgements This research was partially supported by the Research Grants Council of Hong Kong, grant number CUHK4245/04E. The authors greatly appreciate Prof. Xiaoling Sun for his suggestions and careful proofreading of the manuscript of this paper.

Appendix

Table 1 Values of $f(x) = (x - \bar{x})^T Q(x - \bar{x})$ on $\mathbb{X}$

x_2	x_1							
	0	1	2	3	4	5	6	7
0	*3	28	139	334	616	982	1434	1971
1	80	*6	18	115	297	565	918	1356
2	271	99	12	*10	93	262	516	856
3	578	306	120	20	*4	75	230	471
4	999	629	344	144	30	*2	58	200
5	1535	1066	682	384	171	43	**1	44
6	2185	1617	1135	738	426	200	59	*3
7	2951	2284	1703	1207	796	471	231	77

* Discrete local minimum (see Definition 5)

** Discrete global minimum

Table 2 Summary of the computational results for Problem 1

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 5$			$N_{\text{test}} = 5$		
Minimum	6	1.71	4.40	2002	5911
Mean	28.20	2.87	5.52	3547.00	7456.00
Median	36.00	2.66	5.37	3523.00	7432.00
Maximum	43	4.43	6.94	5883	9792

Table 3 Summary of the computational results for Problem 2

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 4$			$N_{\text{test}} = 9$		
Minimum	1	0.83	2.14	869	3011
Mean	3.33	1.31	2.63	1625.78	3767.78
Median	4.00	1.18	2.46	1526.00	3668.00
Maximum	6	1.90	3.22	2707	4849

Table 4 Summary of the computational results for Problem 3

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 2$			$N_{\text{test}} = 7$		
Minimum	0	0.32	34.43	1007	60687
Mean	32.86	4.11	38.04	8521.43	68196.29
Median	2.00	2.56	36.26	7289.00	66969.00
Maximum	113	12.08	45.81	21925	81593

Table 5 Summary of the Computational Results for Problem 4

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 2$			$N_{\text{test}} = 7$		
Minimum	3	0.98	164.64	3671	285475
Mean	5.14	85.40	248.89	163109.43	444887.71
Median	4.00	39.46	203.04	75820.00	357594.00
Maximum	11	182.96	346.41	344828	626602

Table 6 Summary of the Computational Results for Problem 5

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 2$			$N_{\text{test}} = 9$		
Minimum	192	8.30	927.36	18867	1429317
Mean	202.67	312.65	1278.54	526903.33	1937345.33
Median	208.00	449.14	1395.78	720576.00	2131014.00
Maximum	208	482.20	1598.90	848445	2258883

Table 7 Summary of the Computational Results for Problem 6

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 4$			$N_{\text{test}} = 10$		
Minimum	53	2275.61	3493.56	4273411	6560251
Mean	53.00	2408.32	3651.27	4444392.00	6731232.00
Median	53.00	2434.02	3661.15	4441824.00	6728664.00
Maximum	53	2522.89	3770.81	4610237	6897077

Table 8 Summary of the computational results for Problem 7

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 25$			$N_{\text{test}} = 9$		
Minimum	0	0.57	137.80	1351	242659
Mean	1.00	0.85	145.71	2488.89	243796.89
Median	1.00	0.81	145.26	2389.00	243697.00
Maximum	2	1.58	156.19	4527	245835
$n = 50$			$N_{\text{test}} = 9$		
Minimum	0	1.76	1077.36	5201	1920991
Mean	1.11	2.89	1086.75	9707.33	1925497.33
Median	1.00	2.63	1086.36	9460.00	1925250.00
Maximum	2	5.80	1105.35	18434	1934224
$n = 100$			$N_{\text{test}} = 9$		
Minimum	0	5.64	8844.48	20401	15298884
Mean	1.11	10.83	8864.24	37741.56	15316224.56
Median	1.00	9.60	8866.64	36435.00	15314918.00
Maximum	2	23.74	8916.17	74359	15352842

Table 9 Summary of the computational results for Problem 8

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 25$			$N_{\text{test}} = 9$		
Minimum	0	0.49	121.92	1351	218686
Mean	1.00	51.78	176.49	90056.67	305712.11
Median	1.00	0.78	129.83	2470.00	219805.00
Maximum	2	117.18	241.23	206033	419589
$n = 50$			$N_{\text{test}} = 9$		
Minimum	0	1.54	930.14	5201	1707273
Mean	1.00	403.13	1343.91	728415.00	2423846.56
Median	1.00	2.69	956.33	9219.00	1711291.00
Maximum	2	917.55	1857.34	1641022	3328153
$n = 100$			$N_{\text{test}} = 9$		
Minimum	0	5.08	7539.03	20401	13506266
Mean	1.00	3285.58	10845.75	5879746.89	19333796.78
Median	1.00	9.90	7626.02	35944.00	13521809.00
Maximum	2	7473.98	15055.71	13233642	26647923

Table 10 Summary of the Computational Results for Problem 9

	N_{iter}	T_{final}	T_{stop}	N_{final}	N_{stop}
$n = 25$			$N_{\text{test}} = 10$		
Minimum	11	4.85	137.17	10142	255392
Mean	11.60	50.59	183.76	96382.20	341632.20
Median	12.00	43.42	176.85	81236.50	326486.50
Maximum	12	104.56	236.58	202344	447594
$n = 50$			$N_{\text{test}} = 10$		
Minimum	24	17.34	1059.14	40954	1990854
Mean	24.40	117.65	1164.25	234692.30	2184592.30
Median	24.00	20.41	1068.01	50453.50	2000353.50
Maximum	25	437.92	1480.64	843791	2793691
$n = 100$			$N_{\text{test}} = 10$		
Minimum	48	68.86	8616.65	161816	15711616
Mean	49.20	463.98	9029.10	909959.70	16459759.70
Median	49.00	75.06	8655.06	183134.00	15732934.00
Maximum	50	1860.80	10465.85	3482697	19032497

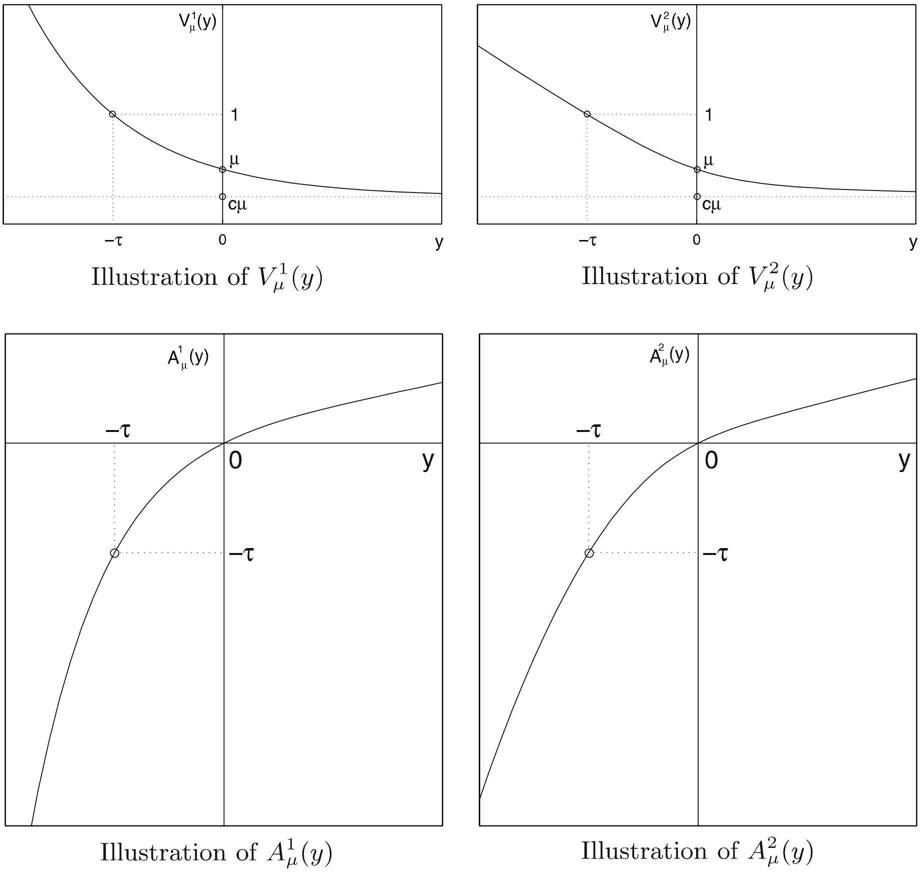

Fig. 1 Illustrations of $V^1_\mu(y)$, $V^2_\mu(y)$, $A^1_\mu(y) = y \cdot V^1_\mu(y)$ and $A^2_\mu(y) = y \cdot V^2_\mu(y)$ in Example 3 with $c = \mu = 0.5$ and $\tau = 1$

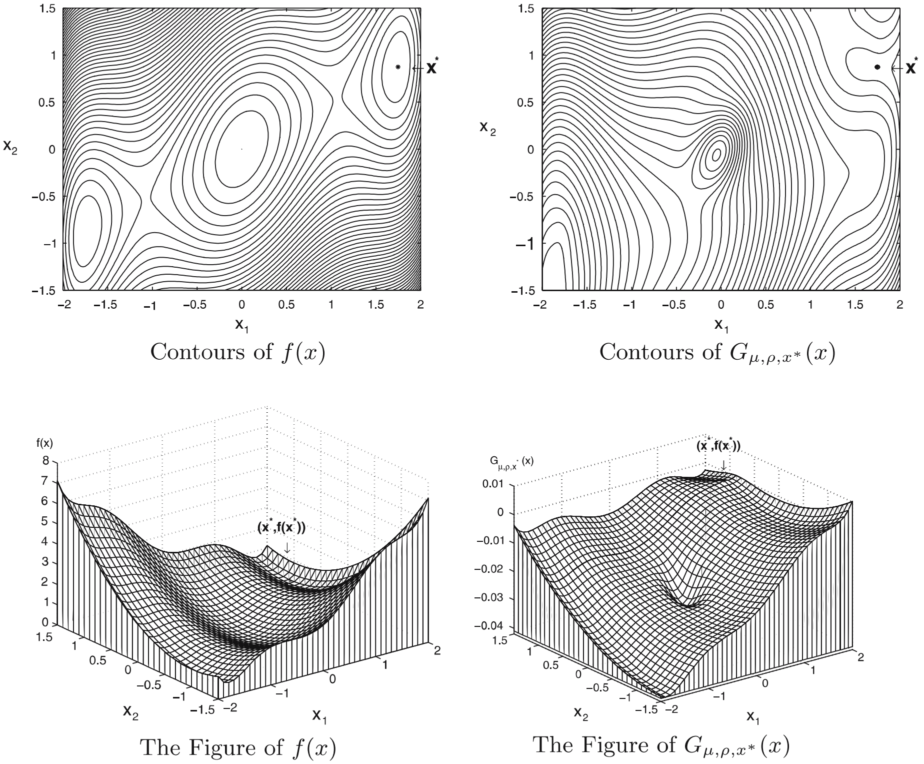

Fig. 2 Illustration of the discrete global descent function for Example 4

References

1. Borghi, C.A., Casadei, D., Febbri, M., Serra, G.: Reduction of the torque ripple in permanent magnet actuators by a multi-objective minimization technique. IEEE Trans. Magnet. **34**(5), 2869–2872 (1998)
2. Borghi, C.A., Fabbri, M.: A combined technique for the global optimization of the inverse electromagnetic problem solution. IEEE Trans. Magnet. **33**(2), 1947–1950 (1997)
3. Borghi, C.A., Febbri, M., Di Barba, P., Savini, A.: A comparative study of loneys solenoid by different techniques of global optimization. Int. J. Appl. Electromagnet. Mech. **10**(5), 417–423 (1999)
4. Conley, W.: Computer Optimization Techniques. Petrocelli Books Inc, New York (1980)
5. Fisher, M.L.: The Lagrangian relaxation method for solving integer programming problems. Manage. Sci. **27**(1), 1–18 (1981)
6. Ge, R.-P.: A filled function method for finding a global minimizer of a function of several variables. Math. Program. **46**(2), 191–204 (1990)
7. Ge, R.-P., Huang, C.-B.: A continuous approach to nonlinear integer programming. Appl. Mathe. Comput. **34**(1), 39–60 (1989)
8. Ge, R.-P., Qin, Y.-F.: A class of filled functions for finding global minimizers of a function of several variables. J. Optim. Theory Appl. **54**(2), 241–252 (1987)
9. Geoffirion, A.M.: Lagrangian relaxation for integer programming. Mathe. Program. Study **2**, 82–114 (1974)
10. Goldstein, A.A., Price, J.F.: On descent from local minima. Mathe. Comput. **25**(115), 569–574 (1971)

11. Gupta, O.K., Ravindran, A.: Branch and bound experiments in convex nonlinear integer programming. Manage. Scie **31**(12), 1533–1546 (1985)
12. Hock, W., Schittkowski, K.: Test Examples for Nonlinear Programming Codes. Springer-Verlag, New York (1981)
13. Körner, F.: A new branching rule for the branch and bound algorithm for solving nonlinear integer programming problems. BIT **28**(3), 701–708 (1988)
14. Li, D., Sun, X.-L.: Success guarantee of dual search in integer programming: *p*th power Lagrangian method. *J. Global Optim.* **18**, 235–254 (2000)
15. Li, D., Sun, X.: Nonlinear Integer Programming. Springer, Boston (2006)
16. Li, D., White, D.J.: *p*th Power Lagrangian method for integer programming. Ann. Operat. Res. **98**, 151–170 (2000)
17. Liu, X.: Finding global minima with a computable filled function. *J. Global Optim.* **19**(2), 151–161 (2001)
18. Liu, X., Xu, W.: A new filled function applied to global optimization. Comp. Operat. Res. **31**(1), 61–80 (2004)
19. Mohan, C., Nguyen, H.T.: A controlled random search technique incorporating the simulated annealing concept for solving integer and mixed integer global optimization problems. Comput. Optim. Appl. **14**, 103–132 (1999)
20. Moré, J.J., Garbow, B.S., Hillstrom, K.E.: Testing unconstrained optimization software. ACM Trans. Math. Software **7**(1), 17–41 (1981)
21. Ng, C.-K.: High performance continuous/discrete global optimization methods. Ph.D. Dissertation, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong S.A.R., P.R.China (2003)
22. Ng, C.-K., Li, D., Zhang, L.-S.: Filled function approaches to nonlinear integer programming: a survey. In: Hou, S.H., Yang, X.M., Chen, G.Y. (eds.) Frontiers in Optimization and Control. Kluwer Academic Publishers. To appear (2005a)
23. Ng, C.-K., Zhang, L.-S., Li, D., Tian, W.-W.: Discrete filled function method for discrete global optimization. Comput. Optim. Appl. **31**(1), 87–115 (2005b)
24. Schittkowski, K.: More Test Examples for Nonlinear Programming Codes. Springer-Verlag (1987)
25. Sun, X.-L., Li, D.: Asymptotic strong duality for bounded integer programming: a logarithmic-exponential dual formulation. Math. Operat. Res. **25**, 625–644 (2000)
26. Xu, Z., Huang, H.-X., Pardalos, P.M., Xu, C.-X.: Filled functions for unconstrained global optimization. J. Global Optim. **20**(1), 49–65 (2001)
27. Yang, X.Q., Goh, C.J.: A nonlinear Lagrangian function for discrete optimization problems. In: Migdalas, A., Pardalos, P.M., and Värbrand, P. (eds.) From Local to Global Optimization. Kluwer Academic Publishers, Dordrecht (2001)
28. Zhang, L.-S., Gao, F., Zhu, W.-X.: Nonlinear integer programming and global optimization. J. Comput. Math. **17**(2), 179–190 (1999)
29. Zhang, L.-S., Ng, C.-K., Li, D., Tian, W.-W.: A new filled function method for global optimization. J. Global Optim. **28**(1), 17–43 (2004)
30. Zheng, Q., Zhuang, D.-M.: Integral global minimization: algorithms, implementations and numerical tests. J. Global Optim **7**(4), 421–454 (1995)
31. Zhu, W.-X.: An approximate algorithm for nonlinear integer programming. Appl. Math. Comput. **93**(2–3), 183–193 (1998)