Front. Eng. Manag. 2018, 5(4): 515–523
https://doi.org/10.15302/J-FEM-2018044

RESEARCH ARTICLE

Panos M. PARDALOS, Mahdi FATHI

A discussion of objective function representation methods in global optimization

Abstract Non-convex optimization can be found in several smart manufacturing systems. This paper presents a short review on global optimization (GO) methods. We examine decomposition techniques and classify GO problems on the basis of objective function representation and decomposition techniques. We then explain Kolmogorov's superposition and its application in GO. Finally, we conclude the paper by exploring the importance of objective function representation in integrated artificial intelligence, optimization, and decision support systems in smart manufacturing and Industry 4.0.

Keywords global optimization, decomposition techniques, multi-objective, DC programming, Kolmogorov's superposition, space-filling curve, smart manufacturing and Industry 4.0

1 Global optimization (GO) methods

Global non-convex programs can be solved using several approaches according to recent advances in GO literature (Pardalos and Rosen, 1986; Pardalos, 1991; Bomze et al., 1997; Pardalos and Wolkowicz, 1998; Horst et al., 2000; Nowak, 2005; Floudas and Pardalos, 2013; Horst and Pardalos, 2013; Floudas and Pardalos, 2014). These approaches can be divided into exact methods that can find and verify global solutions and heuristic methods, which only seek global solutions without checking optimality. Heuristics achieve a critical function in the optimization of large-scale non-convex problems and can be applied to provide upper bounds for global optimum, generate cuts and relaxations, and partition feasible sets. Approximation algorithms are kinds of heuristics, wherein performance guarantee is considered estimated error (Fisher, 1980; Hochbaum et al., 1999; Ausiello et al., 2012; Vazirani, 2013). MIP approximation techniques work by approximating univariate functions to piecewise linear function with a performance guarantee for MINLP method. Goemans and Williamson (1995) solved a quadratic binary program using the MaxCut heuristic as first approximation algorithm.

In GO, an algorithm is called finite if it obtains and verifies a global solution in a finite number of step. The exact methods are finite in finding and verifying solution. Moreover, simplex, active set, and enumeration methods are finite for solving LPs, convex QPs, and bounded integer or concave problems. However, interior point and solution methods for SQP as a nonlinear convex program are not finite.

All GO methods create a rough model of the program for finding global solutions. A GO method is called a *sampling heuristic* if the method uses a crude model based on a finite set of points. The considered regions of interest in *sampling heuristic* methods are bounded set. The distribution of points in this region is usually denser and should consider random behavior to obtain all possible solutions. In the continuous feasible region, the possible random sample is infinite, and a GO solution is not guaranteed. Moreover, the sample can prove that the method converges with probability that is arbitrarily close to 1. A GO method is called a *relaxation-based method* if the method uses relaxation as a crude model, such as a mathematical model, which is easier to solve than the original problem. The crude model influences the problem description. Modeling the problem in an aggregated form is efficient for sampling heuristics with few variables and a simple, feasible set in a disaggregated form for relaxation-based method with objective functions and constraints that can be relaxed.

Relaxation-based heuristics are classified into three relaxation-based methods classes, which include *branch-and-bound methods.* This method divides the GO problem

Received May 2, 2018; accepted August 2, 2018

Panos M. PARDALOS (✉), Mahdi FATHI
Department of Industrial Engineering and Systems Engineering, University of Florida, Gainesville, FL 116595, USA
E-mail: pardalos@ise.ufl.edu

into subproblems based on partitioning of the feasible set. S*uccessive relaxation methods* successively improve an initial relaxation without dividing it into subproblems. *Heuristics* retrieve potential solutions from a given relaxation without modifying the relaxation.

The MINLP solver technology should be further developed, and additional details on GO (Pardalos and Rosen, 1987; Pintér, 1996; Horst et al., 2000; Neumaier, 2004; Schichl, 2010; Horst and Pardalos, 2013; Horst and Tuy, 2013;), MINLP methods (Floudas et al., 1989; Grossmann and Kravanja, 1997; Grossmann, 2002; Tawarmalani and Sahinidis, 2002; Floudas, 2013), and sampling heuristics (Torn and Zilinskas, 1989; Boender and Romeijn, 1995; Strongin and Sergeyev, 2000) should be identified. In summary, GO methods can be classified as follows:

- **Sampling heuristics**: 1) Ultistar (Strongin and Sergeyev, 2000), 2) Clustering method (Becker and Lago, 1970; Dixon and Szegö, 1974; Torn and Zilinskas, 1989), 3) Evolutionary algorithm (Forrest, 1993), 4) Simulated annealing (Metropolis et al., 1953; Kirkpatrick et al., 1983; Locatelli M, 2002), 5) Tabu search (Glover and Laguna, 1997; Mart et al., 2018), 6) Statistical GO (Mockus J, 2012), 7) Greedy randomized adaptive search procedure (Resende and Ribeiro, 2003; Hirsch et al., 2007)
- **Branch-and-bound methods**: 1) Branch-and-bound (Smith and Pantelides, 1996; Vaidyanathan and El-Halwagi, 1996; Smith and Pantelides, 1999; Horst and Tuy, 2013), 2) Branch-and-cut (Padberg and Rinaldi, 1991), 3) Branch-and-reduce (Sahinidis, 1996), 4) Branch-and-price, 5) Branch-cut-and-price, 6) Branch-and-infer (Van Hentenryck et al., 1997; Bliek, 1998; Boddy and Johnson, 2002; Sellmann and Fahle, 2003; Hooker, 2011).
- **Successive approximation method**: 1) Extended cutting-plan method (Westerlund and Pettersson, 1994, 1995; Westerlund et al., 2001), 2) Generalized bender decomposition (Geoffrion, 1972; Floudas et al., 1989; Paules and Floudas, 1989), 3) Outer approximation (Duran and Grossmann, 1986; Kocis and Grossmann, 1987; Viswanathan and Grossmann, 1990; Fletcher and Leyffer, 1994; Zamora and Grossmann, 1998a, 1998b; Grossmann, 2002; Kesavan et al., 2004), 4) Logic-based approach (Türkay and Grossmann, 1996; Vecchietti and Grossmann, 1999), 5) Generalized cross decomposition (Holmberg, 1990), 6) Successive semidefinite relaxation (Lasserre, 2001; Henrion and Lasserre, 2002; Kojima et al., 2003), 7) Lagrangian and domain cut method (Li et al., 2009).
- **Relaxation-based heuristics**: 1) Rounding heuristics (Mawengkang and Murtagh, 1986; Goemans and Williamson, 1995; Burkard et al., 1997; Zwick, 1999), 2) Lagrangian heuristics (Holmberg and Ling, 1997; Nowak and Römisch, 2000), 3) Deformation heuristics (Moré and Wu, 1997; Schelstraete et al., 1999; Alperin and Nowak, 2005), 4) MIP approximation (Neumaier, 2004), 5) Successive linear programming (Palacios-Gomez et al., 1982).

2 Decomposition theory

Large-scale problems can be solved by splitting them into subproblems, which are coupled by a master problem either in parallel or in sequence. The Dantzig–Wolfe decomposition employs separability to decompose a GO problem to subproblems; this method is one of the first decomposition approaches for linear programming that could be optimized in parallel (Dantzig and Wolfe, 1960). This method considers dual problem as a master problem, which coordinates the solutions and iterative modifications of the subproblems. The extension of Dantzig–Wolfe decomposition was applied to the nonlinear convex problem, and the Lagrangian dual is solved by using the cutting plane method. Details regarding decomposition methods in convex and non-convex GO problems are found in (Kelly et al., 1998; Bertsekas, 1999; Horst et al., 2000; Babayev and Bell, 2001; Svanberg, 2002; Palomar and Chiang, 2006; Zhang and Wang, 2006; Boyd et al., 2007; Chiang et al., 2007; Zheng et al., 2013; Rockafellar, 2016; Rahmaniani et al., 2017; Nowak et al., 2018). In general, decomposition techniques can be classified into dual and primal decomposition methods.

2.1 Primal decomposition

The following program with objective function is considered:

$$\left\{max_{y,x_i}\ \sum_i f_i(x_i);\ \text{subject to}: x_i\in X_i\right\},\qquad(1)$$

where $\forall i\ A_i x_i \leqslant y$, and $y\in Y$. Primal decomposition can be applied wherever a coupling variable is set to a fixed value. Thereafter, the GO problem is decoupled into several subproblems for each i as:

$$\{max_{x_i}\ f_i(x_i);\ \text{subject to}: x_i\in X_i,\ A_i x_i\leqslant y\}.\qquad(2)$$

The master problem updates the coupling variable by solving:

$$\left\{max_y\ \sum_i f_i^*\ (y);\ \text{subject to}\ :\ y\in Y\right\},\qquad(3)$$

where $\lambda, f_i^*(y)$ is the optimal objective value in (2). Therefore, Problems (2) and (3) are convex optimization problems if Problem (1) is convex. The gradient method solves Problem (3). Therefore, the optimal Lagrange multiplier, $\lambda_i^*(y)$ in (2), the subgradient for each $f_i^*(y)$ obtained by $s_i(y)=\lambda_i^*(y)$), and Problem (2) can be solved by $\boldsymbol{y}$, where $s(y)=\sum_i s_i(y)=\sum_i \lambda_i^*(y)$ is the global subgradient.

2.2 Dual decomposition

Dual decomposition is suitable when a coupling constraint and its relaxation exist. The GO problem is divided into several subproblems.

$$\left\{max_{\boldsymbol{x}_i}\sum_i f_i(\boldsymbol{x}_i);\ \text{subject to}: \boldsymbol{x}_i \in \boldsymbol{X}_i,\ \forall i \sum_i \boldsymbol{h}_i(\boldsymbol{x}_i) \leqslant \boldsymbol{c}\right\}. \tag{4}$$

The following equation is obtained by applying Lagrangian relaxation to the coupling constraint in Problem (4):

$$\left\{max_{x_i}\sum_i \mathrm{f}_i(x_i)\ -\lambda^{\mathrm{T}}\left(\sum_i h_i(x_i)\text{-}c\right);\right.$$

$$\left.\text{subject to}:\ x_i \in X_i\ \forall i\right\}. \tag{5}$$

The Lagrangian subproblem for each i decouples Problem (5)

$$\{max_{x_i}\ f_i(x_i) - \lambda^{\mathrm{T}}(h_i(x_i)c\ )\ ;\ \text{subject to}:\ x_i \in X_i\}. \tag{6}$$

The dual variables are updated from the master dual problem as follows:

$$\left\{min_\lambda = \sum_i g_i(\lambda)\ + \lambda^{\mathbf{T}}\boldsymbol{c};\ \text{subject to}:\ \lambda \geqslant 0\right\}, \tag{7}$$

where $g_i(\lambda)$ is the dual function obtained as the maximum value of the Lagrangian solved in Problem (6) for a given λ. Thus, a gradient method can solve Problem (7), and the subgradient for each $g_i(\lambda)$ obtained by $s_i(\lambda) = -h_i(x_i^*(\lambda))$, where $_i(\lambda)$ is the optimal solution of Problem (6) for a given λ. The global subgradient is $s(\lambda) = \sum_i s_i(y) + c = c - \sum_i h(x_i^*(\lambda))$. Problem (6) can be independently and locally solved with knowledge of λ.

3 Objective function representation based on decomposition methods

3.1 Separable optimization

The choice of decomposition (of objective function) influences the choice of the algorithm for solving the corresponding mathematical program.

Definition 1: Separable optimization Problem (Horst et al., 2000)

$$\{min_{x\in\Re^n}F_0(x)\ \textit{subject to}: F_i(x)\leqslant b_i,\ l_i\leqslant x_i\leqslant u_i,$$

$$i = 1,\ldots,m\}, \tag{8}$$

where $F_i(x) = \sum_{j=1}^n F_{ij}(x_j),\ i = 0,1,\ldots,m.$

3.2 Factorable optimization

McCormick (1983, 1974, 1976) introduced factorable programming. A factorable program takes the following form

$$\{min_{x\in\Re^n}X^L(x)\ \textit{subject to}: l_i\leqslant X^i(x)\leqslant u_i,\ i = 1,\ldots,L-1\}, \tag{9}$$

where $X^i : \Re^n \mapsto \Re$

$X^i(x) = x_i$ for $i = 1,\ldots,n$ and $X^p(x)$, $p = 1,\ldots,i-1$, function X^i is $X^i(x) = \sum_{p=1}^{i\text{-}1} T_p^i(X^p(x)) + \sum_{p=1}^{i-1}\sum_{q=1}^{p} V_{q,p}^i (X^p(x)).U_{p,q}\ (X^q(x))$, where T's, U's, and V's are the transformation functions of a single variable. The lower and upper bounds $l_i \leqslant u_i$ are given constants. The function $X^i(x), i = 1,\ldots,L$ can be written as factorable functions. McCormick (1974) developed a factorable programming language integrated with SUMT (Mylander et al., 1971) for NLPs. The functions $X^i(x),\ i = 1,\ldots,L$ are called concomitant variable functions (cvfs). The cvfs includes separable and quadratic terms.

3.3 Almost block separable optimization

The following problem is considered:

$$min_{x\in R^n}f(x)\ = f_1(u,y) + f_2(v,y), \tag{10}$$

where $x = (u,v,y) \in \Re^n$ *and* $u \in \Re^{n1}, v \in \Re^{n2}, y \in \Re^{n3}$, $n_1 + n_2 + n_3 = n$, and y are called complicated variables [usually $n_1,n_2 \gg n_3$]

Let $\varphi_1(y)\ = min_u f_1(u,y), \varphi_2(y)\ = min_v f_2(v,y)$. The problem is equivalent to:

$$min_y \varphi_1(y)\ + \varphi_2(y). \tag{11}$$

If f_1 and f_2 are convex, then $\varphi_1(y)$ and $\varphi_2(y)$ are convex.

3.4 DC optimization problems

3.4.1 Continuous DC programming

One of the special non-convex programs is DC programming. DC function and dual DC programming are defined as follows:

Definition 2: DC function (Horst et al., 2000; Wu et al., 2018)

A real-valued function $f : \Re^n \mapsto \Re \cup \{+\infty, -\infty\}$ subject to:

$$\{f(x)\ = g(x)\ - h(x),\ \forall x \in \Re^n\}, \tag{12}$$

where $g,\ h : \Re^n \mapsto \Re\ \cup +\infty$ is a convex function and is a

DC function for any h and g.

Definition 3: DC program (Horst et al., 2000; Wu et al., 2018)

The following model is called a DC program

$$\{min\ f_0(x)\ \text{subject to}: f_i(x) \leqslant 0,\ \forall i = 1,2...,n\},\quad (13)$$

if $f_i(x)$ are DC functions ($i = 0,1,2,...,n$) and it is the same as the following DC program. Then,

$$inf_{x_\in <\Re^n} f(x) \ = g(x) - h(x).\quad (14)$$

Hartman Theorem 1. The following DC programs are equal:

$$\begin{cases} sup\ f(x) : x \in C, f,\ C:\ \text{convex} \\ \inf\ g(x) - h(x) : x \in \Re^n,\ g,\ h:\ \text{convex} \\ \inf\ g(x) - h(x) : x \in C,\ f_1(x) - f_2(x) \\ \leqslant 0,\ g,\ h,\ f_1,\ f_2,\ C: \text{all convex} \end{cases}$$

Hartman Theorem 2. A function f is locally DC if an $\in$ -ball on which DC exists. Every function that is locally DC is considered a DC proposition. Let f_i be DC functions for $i = 1,...,m$. Thus, $\left\{\sum_i \lambda_i f_i(x) \text{ for } \lambda_i \in \Re\right\}$; $\{max_i f_i(x)\}$; $\{min_i f_i(x)\}$; $\{\Pi_i f_i(x)\}$; and $\{f_i\}$ are twice continuously differentiable DC. Moreover, (gof) is DC if f is DC and g is convex, and every continuous function on C (convex set) is the limit of a sequence of uniformly converging DC functions.

Definition 4: Subgradient of convex function (Horst et al., 2000; Wu et al., 2018)

A vector x^* is a subgradient of a convex function h at a point x if $h(z) \geqslant h(x) + \langle x^*, z - x\rangle$, where $\langle x,y\rangle = \sum_{i=1}^n x_i y_i$ is the inner product of two vectors with the same dimension. The subdifferential of $h(x)$ is the set of all subgradients.

Definition 5: Conjugate functions (Horst et al., 2000; Wu et al., 2018)

A conjugate function $h^*: \Re^n \mapsto \Re \cup +\infty$ of a convex function $h: \Re^n \mapsto \Re \cup +\infty$ is:

$$h^*(p) := sup_{y \in \Re^n} \langle y,x\rangle - h(x).\quad (15)$$

Theorem 3: The conjugate function $h^*(y)$ of $h(x)$ is convex. If $h(x)$ is a closed proper convex function, then the bi-conjugate of h is itself, that is, $h^{**} = h$.

Theorem 4 (Toland–Singer duality): Given closed convex functions $g,\ h:\ \Re^n \mapsto \Re \cup +\infty$, then:

$$inf_{x \in \Re^n}\{g(x) - h(x)\} = inf_{p \in \Re^n}\{h^*(p) - g^*(p)\}.\quad (16)$$

Definition 6: DC algorithm (Horst et al., 2000; Wu et al., 2018)

The following algorithm is used for obtaining a local optimal solution for the DC program.

Step 0 Find an inintial solution $x^0 \in dom_R(g)$. Set $t := 0$.

Step 1 Find $p^t \in \partial_R h(x^t)$.

Step 2 Find $p^{t+1} \in \partial_R g^*(p^t)$, where g^* is the "conjugate" of g.

Step 3 If $f(x^{t+1}) = f(x^t)$, stop. Otherwise, set $t := t+1$, go to Step 1.

where $x^{t+1} = argmin_{y \in R^n}\{g(y) - h(x^t) - \langle p, y - x^t\rangle\}$.

3.4.2 Continuous relaxations for discrete DC programming

The positive support of $x \in Z^n$ is presented as follows: $supp^+(x) := \{i \in \{1,2,...,n\} : x_i > 0\}$.

The indicator vector χ_S is defined by:

$$\chi_S(i) = \begin{Bmatrix} 1 & i \in S \\ 0 & i \notin S \end{Bmatrix}.$$

$M^\natural$-convex and $L^\natural$-convex are two common discrete functions:

1) $M^\natural$-convex functions are defined as $\forall x,y \in \mathbb{Z}^n and\ i \in supp^+(x - y)$, function h: $\mathbb{Z}^n \mapsto \mathbb{Z}\ +\infty$ is $M^\natural$-convex if it satisfies:

$$h(x) + h(y) \geqslant min\{h(x - \chi_i) + h(x + \chi_i)\},\quad (17)$$

$$min_{j \in supp^+(x-y)} h(x - \chi_i + \chi_j) + h(y + \chi_i - \chi_j).\quad (18)$$

2) $L^\natural$-convex functions are defined as $\forall x,\ y \in \mathbb{Z}^n$, $h: \mathbb{Z}^n \mapsto \mathbb{Z} \cup +\infty$ is $L^\natural$ -convex if it satisfies:

$$h(x) + h(y) \geqslant h\left(\left\lceil\frac{x+y}{2}\right\rceil\right) + h\left(\left\lfloor\frac{x+y}{2}\right\rfloor\right).\quad (19)$$

Consider the following discrete DC program:

$$\{Inf\ f(x) = g(x) - h(x)\ \text{subject to}: x \in \mathbb{Z}^n\}.\quad (20)$$

The four kinds of discrete DC programs include $M^\natural - L^\natural$, $M^\natural - M^\natural$, $L^\natural - L^\natural$, and $L^\natural - M^\natural$, wherein the first three are NP-hard, and the last one on $\{0,1\}^n$ is in P, can be defined on the basis of $M^\natural$ and $L^\natural$-convex function definitions (Kobayashi, 2014; Maehara et al., 2018) .

We assume functions $g,h : \mathbb{Z}^n \mapsto \Re \cup \{+\infty\}$. The effective domain of g is $dom_Z g := \{x \in \mathbb{Z}^n : g(x) < +\infty\}$.

The convex closure $\overline{g}(x) : \Re^n \mapsto \Re \cup \{+\infty\}$ of g is:

$$\overline{g}(x) = sup\{s(x) : s \text{ is an affine function},$$

$$s(y) \leqslant g(y) (y \in \mathbb{Z}^n)\}.\quad (21)$$

A convex extension $\hat{g}: \Re^n \mapsto \Re \cup \{+\infty\}$ of g is a convex function with the same function value on $x \in \mathbb{Z}^n$.

We assume

$\tilde{f}(x) := \overline{g}(x) - \hat{h}(x)$. Then $\tilde{f}(x) := g(x) - h(x),\ \forall x \in \mathbb{Z}^n$.

Thus:

$$inf_{x \in \mathbb{Z}^n}\{g(x) - h(x)\} = inf_{x \in \mathbb{Z}^n} \tilde{f}(x) \geqslant inf_{x \in \Re^n} \tilde{f}(x).\quad (22)$$

Theorem 5: For convex extensible functions g, h : $\mathbb{Z}^n \mapsto \Re \cup +\infty$ with $dom_Z g$ bounded and $dom_z g \in dom_Z h$:

$$inf_{z \in \mathbb{Z}^n}\{g(z)-h(z)\} = inf_{x \in \Re^n}\left\{\overline{g}(x)-\hat{h}(x)\right\}, \quad (23)$$

where $\overline{g}(x)$ is the linear closure of $g(x)$, and $\hat{h}(x)$ is any convex extension of $h(x)$.

We found that the discrete DC programming (20) is equivalent to the corresponding continuous relaxation DC programming based on Theorem 5.

3.5 DI optimization problems

Total and partial monotonicity are related to monotonicity for all and some variables with many GO applications. The d.i. monotonic optimization with increasing functions in $\Re^n_+$ can be generally described as follows:

$$\{min\ f(x)-g(x)\ \text{subject to}: f_i(x)-g_i(x) \leqslant 0,\ i=1,...,m\}. \quad (24)$$

Let $g(x) = 0$, and,

$$\forall i, f_i(x)-g_i(x) \leqslant 0 \Leftrightarrow max_{1 \leqslant i \leqslant m}\{f_i(x)-g_i(x)\} \leqslant 0$$

$$\Leftrightarrow F(x)-G(x) \leqslant 0, \quad (25)$$

with increasing F, and G $\left(F(x) = max_i\{f_i(x)+\sum_{i \neq j} g_j(x)\},\ G(x) = \sum_i g_i(x)\right)$.

Then, the problem is reduced to:

$$\{minf(x);\ \text{subject to}:\ F(x)+t \leqslant F(b),\ G(x)+t \geqslant F(b),$$

$$0 \leqslant t \leqslant F(b)-F(0),\ x \in [0,b] \subset \Re^n_+\}. \quad (26)$$

For any x, x' where $x' \leqslant x$, *if* $x \in G$, then $x' \subseteq G$, a set $G \subseteq \Re^n_+$ is normal.

Many GO problems, including polynomial, multiplicative, Lipschitz optimization problems, and non-convex quadratic programming, can be considered monotonic optimization problems.

3.6 Decomposition and multi-objective optimization

We consider the following problems:

$$P1 : min_{x \in D \subseteq \Re^n}\ F(x) = f_1(x) + ... + f_k(x), \quad (27)$$

$$P2 : min_{x \in D \subseteq \Re^n}\ f(x) = (f_1(x),...,f_k(x)). \quad (28)$$

Objective function $F(x)$ in many GO problems can be represented by the summation of k relatively simple functions as $F(x) = f_1(x) + f_2(x) + ... + f_k(x)$. P2 is a multi-objective optimization problem. Let $E(f,D) \subseteq D$ be the set of all Pareto optimal solutions in D. We obtain the following theorems for optimal solutions of P1 and the optimal Pareto frontier of P2.

Theorem 6: If $\overline{x}$ is an optimal solution of P1, then $x \in E(f, D)$ of P2.

Theorem 7: Let $h_i(t)$ be a monotonic increasing function for $i = 1,...,k$. We consider the multi-objective optimization problem $min_{x \in D \subseteq \Re^n} h(x) = (h_1(f_1(x)),...,h_m(f_k(x)))$. Then, $E(f, D) = E(h, D)$. (Miettinen, 1999; Chinchuluun and Pardalos, 2007; Pardalos et al., 2008; Du and Pardalos, 2013; Migdalas et al., 2013; Pardalos et al., 2017)

Theorems 1 and 2 show that the extended Pareto optimal frontier set $E(h,D)$ can be obtained by solving P2 and searching for the optimal $\overline{x}$ of P1 from $E(h,D)$.

P2 can be a multi-objective optimization problem (MaOP). The algorithms for solving MaOPs can be classified as: 1) **Algorithm adaptation methods**, which modify/extend the classical EMO algorithms for solving MaOPs, including preference-based MOEA (PICEA; PBEA), Pareto-based MOEA (NSGA-II; SPEA2), indicator-based MOEA (HypE; SMSEMOA), decomposition-based MOEA (MOEA/D; M2M); and 2) **Problem transformation methods**, which transform the MaOP into a problem with few objectives, including objective selection (σ-MOSS; k-EMOSS; L-PCA) and objective extraction (Gu, 2016) . Refer to Gu (2016) and Mane and Rao (2017), for a review of solution algorithms and real-world applications of MaOPs, such as flight control system, engineering design, data mining, nurse scheduling, car controller optimization, and water supply portfolio planning.

MOEA/D is a mostly used method for solving P2. Its goals can be categorized as: 1) convergence to detect solutions close to the Pareto frontier; 2) diversity to determine well-distributed solutions; and 3) coverage to cover the entire Pareto frontier. Several MOEAs for these goals are found in literature, which can be broadly categorized under three categories, namely, 1) domination-, 2) indicator-, and 3) decomposition-based frameworks (Ehrgott and Gandibleux, 2000; Trivedi et al., 2017).

In MOEA/D literature, three decomposition methods, including the weighted sum (WS), the weighted Tchebycheff (TCH), and penalty based boundary intersection (PBI) approaches.

The ith subproblem of the WS approach is given as:

$$min\ g^{ws}(x|\lambda_i) = \sum_{j=1}^{m} \lambda_i^j f_j(x). \quad (29)$$

This method is efficient for solving convex Pareto solutions with min objective function.

The ith subproblem of the TCH approach is defined as follows:

$$min\ g^{te}(x|\lambda_i, z^*) = max_{1 \leqslant j \leqslant m}\{\lambda_i^j |f_j(x)-z_j^*|\}, \quad (30)$$

where $z^* = (z_1^*,...,z_m^*)^{\mathrm{T}}$ is the ideal reference point with $z_j^* < min\{f_j(x)|x \in \Omega\}$ for $j = 1,2,...,m$.

The ith subproblem of the PBI approach is defined as follows:

$$min\ g^{pbi}(x|\lambda_i,z^*) \ = d_1 + \theta d_2, \tag{31}$$

where $d_1 = \dfrac{//((F(x)-z^*)^{\mathrm{T}}\lambda_i//}{//\lambda_i//}$ and $d_2 = \left\| F(x) - \left(z^* - d_1\dfrac{\lambda_i}{//\lambda_i//}\right)\right\|$. z^* is the reference point shown in (32), and θ is a penalty parameter that should be tuned properly.

4 Kolmogorov's superposition

Kolmogorov (1956) presents the following theorem as Kolmogorov's superposition:

Theorem 8: Continuous real functions $\psi^{p,q}(x)$ (for any integer $n \geqslant 2$) on the closed unit interval $E^1 = [0,1]$ exists similar to continuous real function $f(x_1,...,x_n)$ on the n-dimensional unit cube E^n, which can be shown as:

$$f(x_1,...,x_n) \ = \sum_{q=1}^{2n+1} \chi_q \left[\sum_{p=1}^{n} \psi^{pq}(x_p)\right], \tag{32}$$

$$y^q = \sum_{p=1}^{n} \psi^{pq}(x_p), \tag{33}$$

where $\chi_q(y)$ is a continuous real function (refer to (Arnol'd, 1959; Tikhomirov, 1991) for a brief proof of the theorem). The following equation is obtained for $n = 3$, by setting, $\varphi_q(x_1,\ x_2) = \psi^{1q}(x_1) + \psi^{2q}(x_2)$ and $h_q(y,\ x_3) = \chi_q y + \psi^{3q}(x_3)$: $f(x_1,\ x_2,\ x_3) = \sum_{q=1}^{7}[\varphi_q(x_1,\ x_2),\ x_3]$.

The application of Kolmogorov theorem in GO in space-filling curve is an example of its efficient optimizing functions based on their projection from n dimensions to one dimension (Goertzel, 1999; Lera and Sergeyev, 2010; Sergeyev et al., 2013). Sprecher (Sprecher and Draghici, 2002; Sprecher, 2013; Sprecher, 2014) explored the link between the aforementioned theorem and the space-filling curves from computational algorithms for real-valued continuous functions.

5 Conclusions

This paper reviewed different GO and decomposition methods on the basis of objective function representation. Many GO methods are derived from the branch and bound method, which are inefficient for finding a remarkable solution. This paper provides opportunity for additional research on decomposition techniques based on objective function representation, multi-objective optimization, and Kolmogorov's superposition. The development of other parallel decomposition-based GO methods based on the objective function representation for MINLP, such as Decogo solver (Nowak et al., 2018), can be a challenging area in MINLP solver development. Kolmogorov theorem in GO will be discussed in future studies.

Industry 4.0 is known as the future of smart manufacturing and industrial revolution. Making decentralized decision is critical in Industry 4.0 (Marques et al., 2017). Horizontal and vertical integrations are two principal characteristics in Industry 4.0. Decentralized decision support systems are needed depending on the different types of decisions, including operational, tactical, real-time, and strategic. Many optimization problems are integrated with artificial intelligence in Industry 4.0, in which decision makers (DMs) should make a decentralized decision. This paper will help DMs in Industry 4.0 represent their objective function based on different GO techniques, such as Kolmogorov's superposition and DC programming, which can be solved separately. Finally, Khakifirooz, Pardalos, et al. (2018) and Khakifirooz, Chien, et al. (2018) reported that applications of non-convex optimization in decision support system development for smart manufacturing and Industry 4.0 can be a challenging direction for future research.

Acknowledgements Professor Pardalos' research is partially supported by the Paul and Heidi Brown Preeminent Professorship at ISE, University of Florida. Dr. Mahdi Fathi would like to thank Prof. Murray Brown and Mrs. Helen Brown for their encouragement and support during this research.

References

Alperin H, Nowak I (2005). Lagrangian smoothing heuristics for max-cut. Journal of Heuristics, 11(5–6): 447–463

Arnol'd V I (1959). On the representation of continuous functions of three variables by superpositions of continuous functions of two variables. Matematicheskii Sbornik, 90(1): 3–74

Ausiello G, Crescenzi P, Gambosi G, Kann V, Marchetti-Spaccamela A, Protasi M (2012). Complexity and Approximation: Combinatorial Optimization Problems and Their Approximability Properties. Berlin: Springer Science & Business Media

Babayev D A, Bell G I (2001). An optimization problem with a separable non-convex objective function and a linear constraint. Journal of Heuristics, 7(2): 169–184

Becker R W, Lago G (1970). A global optimization algorithm. In: Proceedings of the 8th Allerton Conference on Circuits and Systems Theory. 3–12

Bertsekas D P (1999). Nonlinear Programming. Belmont: Athena Scientific

Bliek C (1998). Coconut deliverable d1-algorithms for solving nonlinear and constrained optimization problems. The COCONUT Project

Boddy M S, Johnson D P (2002). A new method for the global solution of large systems of continuous constraints. In: Bliek C, Jermann C, Neumaier A, eds. International Workshop on Global Optimization and Constraint Satisfaction. Berlin: Springer

Boender C G E, Romeijn H E (1995). Stochastic methods. In: Pardalos P M, Romeijin H E, eds. Handbook of global optimization. Berlin: Springer

Bomze I M, Csendes T, Horst R, Pardalos P M (1997). Developments in Global Optimization. Berlin: Springer Science & Business Media

Boyd S, Xiao L, Mutapcic A, Mattingley J (2007). Notes on Decomposition Methods. Stanford: Stanford University

Burkard R E, Kocher M, Rüdolf R (1997). Rounding strategies for mixed integer programs arising from chemical production planning. Yugoslav Journal of Operations Research

Chiang M, Low S H, Calderbank A R, Doyle J C (2007). Layering as optimization decomposition: A mathematical theory of network architectures. Proceedings of the IEEE, 95(1): 255–312

Chinchuluun A, Pardalos P M (2007). A survey of recent developments in multiobjective optimization. Annals of Operations Research, 154 (1): 29–50

Dantzig G B, Wolfe P (1960). Decomposition principle for linear programs. Operations Research, 8(1): 101–111

Dixon L C W, Szegö G P (1974). Towards global optimisation. In: Proceedings of a workshop at the University of Cagliari, Italy

Du D Z, Pardalos P M (2013). Handbook of Combinatorial Optimization: Supplement, Vol. 1. Berlin: Springer Science & Business Media

Duran M A, Grossmann I E (1986). An outer-approximation algorithm for a class of mixed-integer nonlinear programs. Mathematical Programming, 36(3): 307–339

Ehrgott M, Gandibleux X (2000). A survey and annotated bibliography of multiobjective combinatorial optimization. OR-Spektrum, 22(4): 425–460

Fisher M L (1980). Worst-case analysis of heuristic algorithms. Management Science, 26(1): 1–17

Fletcher R, Leyffer S (1994). Solving mixed integer nonlinear programs by outer approximation. Mathematical Programming, 66(1–3): 327–349

Floudas C, Aggarwal A, Ciric A (1989). Global optimum search for nonconvex NLP and MINLP problems. Computers & Chemical Engineering, 13(10): 1117–1132

Floudas C A (2013). Deterministic Global Optimization: Theory, Methods and Applications, Vol. 37. Berlin: Springer Science & Business Media

Floudas C A, Pardalos P M (2013). State of the Art in Global Optimization: Computational Methods and Applications, Vol. 7. Berlin: Springer Science & Business Media

Floudas C A, Pardalos P M (2014). Recent Advances in Global Optimization. Princeton: Princeton University Press

Forrest S (1993). Genetic algorithms: Principles of natural selection applied to computation. Science, 261(5123): 872–878

Geoffrion A M (1972). Generalized benders decomposition. Journal of Optimization Theory and Applications, 10(4): 237–260

Glover F, Laguna M (1997). Tabu Search. Berlin: Springer

Goemans M X, Williamson D P (1995). Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. Journal of the Association for Computing Machinery, 42(6): 1115–1145

Goertzel B (1999). Global optimization with space-filling curves. Applied Mathematics Letters, 12(8): 133–135

Grossmann I E (2002). Review of nonlinear mixed-integer and disjunctive programming techniques. Optimization and Engineering, 3(3): 227–252

Grossmann I E, Kravanja Z (1997). Mixed-integer nonlinear programming: A survey of algorithms and applications. In: Biegler L T, Colleman T F, Conn A R, Samtosa F N, eds. Large-scale Optimization with Applications. Berlin: Springer

Gu F Q (2016). Many objective optimization: Objective reduction and weight design. Dissertation for the Doctoral Degree. HongKong: HKBU

Henrion D, Lasserre J B (2002). Solving global optimization problems over polynomials with gloptipoly 2.1. In: Proceedings of International Workshop on Global Optimization and Constraint Satisfaction. Berlin: Springer

Hirsch M J, Meneses C, Pardalos P M, Resende M G (2007). Global optimization by continuous grasp. Optimization Letters, 1(2): 201–212

Hochbaum D, Jansen K, Rolim J D, Sinclair A (1999). Randomization, Approximation, and Combinatorial Optimization. Algorithms and Techniques: In: Proceedings of 3rd International Workshop on Randomization and Approximation Techniques in Computer Science, and 2nd International Workshop on Approximation Algorithms for Combinatorial Optimization Problems. Berlin: Springer Science & Business Media

Holmberg K (1990). On the convergence of cross decomposition. Mathematical Programming, 47(1–3): 269–296

Holmberg K, Ling J (1997). A Lagrangian heuristic for the facility location problem with staircase costs. European Journal of Operational Research, 97(1): 63–74

Hooker J (2011). Logic-Based Methods for Optimization: Combining Optimization and Constraint Satisfaction, Vol. 2. Hoboken: John Wiley & Sons

Horst R, Pardalos P M (2013). Handbook of Global Optimization, Vol. 2. Berlin: Springer Science & Business Media

Horst R, Pardalos P M, Van Thoai N (2000). Introduction to Global Optimization. Berlin: Springer Science & Business Media

Horst R, Tuy H (2013). Global Optimization: Deterministic Approaches. Berlin: Springer Science & Business Media

Kelly F P, Maulloo A K, Tan D K (1998). Rate control for communication networks: Shadow prices, proportional fairness and stability. Journal of the Operational Research Society, 49(3): 237–252

Kesavan P, Allgor R J, Gatzke E P, Barton P I (2004). Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. Mathematical Programming, 100(3): 517–535

Khakifirooz M, Chien C-F, Pardalos F M, Panos M (2018). Management Suggestions on Semiconductor Manufacturing Engineering: An Operations Research and Data Science Perspective. Berlin: Springer

Khakifirooz M, Pardalos P M, Fathi M, Power D J (2018). Decision support for smart manufacturing. Encyclopedia of IST, 5th ed, IGI Global Book

Kirkpatrick S, Gelatt C D Jr, Vecchi M P (1983). Optimization by simulated annealing. Science, 220(4598): 671–680

Kobayashi Y (2014). The complexity of maximizing the difference of two matroid rank functions, METR2014–42. University of Tokyo

Kocis G R, Grossmann I E (1987). Relaxation strategy for the structural optimization of process flow sheets. Industrial & Engineering Chemistry Research, 26(9): 1869–1880

Kojima M, Kim S, Waki H (2003). A general framework for convex relaxation of polynomial optimization problems over cones. Journal of the Operations Research Society of Japan, 46(2): 125–144

Kolmogorov A (1956). On the Representation of Continuous Functions of Several Variables as Superpositions of Functions of Smaller Number of Variables. Berlin: Springer

Lasserre J B (2001). Global optimization with polynomials and the problem of moments. SIAM Journal on Optimization, 11(3): 796–817

Lera D, Sergeyev Y D (2010). Lipschitz and Hölder global optimization using space-filling curves. Applied Numerical Mathematics, 60(1–2): 115–129

Li D, Sun X, Wang J, McKinnon K I (2009). Convergent Lagrangian and domain cut method for nonlinear knapsack problems. Computational Optimization and Applications, 42(1): 67–104

Locatelli M (2002). Simulated annealing algorithms for continuous global optimization. In: Horst R, Pardalos P M, eds. Handbook of Global Optimization. Berlin: Springer

Maehara T, Marumo N, Murota K (2018). Continuous relaxation for discrete DC programming. Mathematical Programming, 169(1): 199–219

Mane S U, Rao M N (2017). Many-objective optimization: Problems and evolutionary algorithms–A short review. International Journal of Applied Engineering Research, 12(20): 9774–9793

Marques M, Agostinho C, Zacharewicz G, Jardim-Goncalves R (2017). Decentralized decision support for intelligent manufacturing in industry 4.0. Journal of Ambient Intelligence and Smart Environments, 9(3): 299–313

Mart R, Panos P, Resende M (2018). Handbook of Heuristics. Berlin: Springer

Mawengkang H, Murtagh B (1986). Solving nonlinear integer programs with large-scale optimization software. Annals of Operations Research, 5(2): 425–437

McCormick G P (1974). A mini-manual for Use of the Sumt Computer Program and the Factorable Programming Language. Stanford: Stanford University

McCormick G P (1976). Computability of global solutions to factorable nonconvex programs: Part iconvex underestimating problems. Mathematical Programming, 10(1): 147–175

McCormick G P (1983). Nonlinear Programming: Theory, Algorithms, and Applications. New York: Wiley

Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H, Teller E (1953). Equation of state calculations by fast computing machines. Journal of Chemical Physics, 21(6): 1087–1092

Miettinen K (1999). Nonlinear Multiobjective Optimization. International Series in Operations Research and Management Science. Berlin: Springer

Migdalas A, Pardalos P M, Värbrand P (2013). Multilevel Optimization: Algorithms and Applications, Vol. 20. Berlin: Springer Science & Business Media

Mockus J (2012). Bayesian Approach to Global Optimization: Theory and Applications, Vol. 37. Berlin: Springer Science & Business Media

Moré J J, Wu Z (1997). Global continuation for distance geometry problems. SIAM Journal on Optimization, 7(3): 814–836

Mylander W C, Holmes R L, McCormick G P (1971). A guide to sumt-version 4: The computer program implementing the sequential unconstrained minimization technique for nonlinear programming (Technical Report RAC-P-63). Mclean: Research Analysis Corporation

Neumaier A (2004). Complete search in continuous global optimization and constraint satisfaction. Acta Numerica, 13: 271–369

Nowak I (2005). Relaxation and decomposition methods for mixed integer nonlinear programming, Vol. 152. Berlin: Springer Science & Business Media

Nowak I, Breitfeld N, Hendrix E M, Njacheun-Njanzoua G (2018). Decomposition-based inner-and outerrefinement algorithms for global optimization. Journal of Global Optimization, (4–5): 1–17

Nowak M P, Römisch W (2000). Stochastic lagrangian relaxation applied to power scheduling in a hydrothermal system under uncertainty. Annals of Operations Research, 100(1–4): 251–272

Padberg M, Rinaldi G (1991). A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems. SIAM Review, 33(1): 60–100

Palacios-Gomez F, Lasdon L, Engquist M (1982). Nonlinear optimization by successive linear programming. Management Science, 28 (10): 1106–1120

Palomar D P, Chiang M (2006). A tutorial on decomposition methods for network utility maximization. IEEE Journal on Selected Areas in Communications, 24(8): 1439–1451

Pardalos P M (1991). Global optimization algorithms for linearly constrained indefinite quadratic problems. Computers & Mathematics with Applications (Oxford, England), 21(6–7): 87–97

Pardalos P M, Migdalas A, Pitsoulis L (2008). Pareto optimality, game theory and equilibria, Vol. 17. Berlin: Springer Science & Business Media

Pardalos P M, Rosen J B (1986). Methods for global concave minimization: A bibliographic survey. SIAM Review, 28(3): 367–379

Pardalos P M, Rosen J B (1987). Constrained Global Optimization: Algorithms and Applications. New York: Springer-Verlag

Pardalos P M, Wolkowicz H (1998). Topics in semidefinite and interior-point methods. American Mathematical Society

Pardalos P M, Zilinskas A, Zilinskas J (2017). Non-convex multi-objective optimization, Vol. 123. Berlin: Springer

Paules G E I V IV, Floudas C A (1989). Apros: Algorithmic development methodology for discrete-continuous optimization problems. Operations Research, 37(6): 902–915

Pintér J D (1996). Global Optimization in Action. Dordrecht: Kluwer Academic Publishers

Rahmaniani R, Crainic T G, Gendreau M, Rei W (2017). The benders decomposition algorithm: A literature review. European Journal of Operational Research, 259(3): 801–817

Resende M G C, Ribeiro C C (2003). Greedy randomized adaptive search procedures. In: Glover F, Kochenberger G, eds. Hand Book of Metaheuristics. Dordrecht: Kluwer Academic Publishers

Rockafellar R T (2016). Problem decomposition in block-separable convex optimization: Ideas old and new. In: Proceedings of the 5th Asian Conference on Nonlinear Analysis and Optimization, Niigata, Japan

Sahinidis N V (1996). Baron: A general purpose global optimization software package. Journal of Global Optimization, 8(2): 201–205

Schelstraete S, Schepens W, Verschelde H (1999). Energy minimization by smoothing techniques: A survey. Molecular Dynamics: from Classical to Quantum Methods

Schichl H (2010). Mathematical Modeling and Global Optimization. Cambridge: Cambridge University Press

Sellmann M, Fahle T (2003). Constraint programming based Lagrangian relaxation for the automatic recording problem. Annals of Operations Research, 118(1–4): 17–33

Sergeyev Y D, Strongin R G, Lera D (2013). Introduction to Global Optimization Exploiting Space-Filling Curves. Berlin: Springer Science & Business Media

Smith E M, Pantelides C C (1996). Global optimisation of general process models. In: Grossmann I E, eds. Global Optimization in Engineering Design. Berlin: Springer

Smith E M, Pantelides C C (1999). A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of non-convex minlps. Computers & Chemical Engineering, 23(4–5): 457–478

Sprecher D (2013). Kolmogorov superpositions: A new computational algorithm. In: Igelnik B, eds. Efficiency and Scalability Methods for Computational Intellect. New York: IGI Global

Sprecher D (2014). On computational algorithms for real-valued continuous functions of several variables. Neural Networks, 59: 16–22

Sprecher D A, Draghici S (2002). Space-filling curves and Kolmogorov superposition-based neural networks. Neural Networks, 15(1): 57–67

Strongin R, Sergeyev Y D (2000). Global Optimization with Non-Convex Constraints. Dordrecht: Kluwer Academic Publishers

Svanberg K (2002). A class of globally convergent optimization methods based on conservative convex separable approximations. SIAM Journal on Optimization, 12(2): 555–573

Tawarmalani M, Sahinidis N V (2002). Convexification and Global Optimization in Continuous and Mixedinteger Nonlinear Programming: Theory, Algorithms, Software, and Applications, Vol. 65. Berlin: Springer Science & Business Media

Tikhomirov V (1991). On the representation of continuous functions of several variables as superpositions of continuous functions of one variable and addition. In: Kolmogorov A N, Shiryaeu A, eds. Selected Works of AN Kolmogorov. Berlin: Springer

Torn A, Zilinskas A (1989). Global Optimization. New York: Springer-Verlag

Trivedi A, Srinivasan D, Sanyal K, Ghosh A (2017). A survey of multiobjective evolutionary algorithms based on decomposition. IEEE Transactions on Evolutionary Computation, 21(3): 440–462

Türkay M, Grossmann I E (1996). Logic-based minlp algorithms for the optimal synthesis of process networks. Computers & Chemical Engineering, 20(8): 959–978

Vaidyanathan R, El-Halwagi M (1996). Global optimization of nonconvex minlps by interval analysis. In: Grossmann I E, eds. Global Optimization in Engineering Design. Berlin: Springer

Van Hentenryck P, Michel L, Deville Y (1997). Numerica: A Modeling Language for Global Optimization. Boston: MIT Press

Vazirani V V (2013). Approximation Algorithms. Berlin: Springer Science & Business Media

Vecchietti A, Grossmann I E (1999). Logmip: A disjunctive 0–1 non-linear optimizer for process system models. Computers & Chemical Engineering, 23(4–5): 555–565

Viswanathan J, Grossmann I E (1990). A combined penalty function and outer-approximation method for minlp optimization. Computers & Chemical Engineering, 14(7): 769–782

Westerlund T, Lundqvist K (2001). Alpha-ECP, version 5.01: An interactive MINLP-solver based on the extended cutting plane method

Westerlund T, Pettersson F (1995). An extended cutting plane method for solving convex minlp problems. Computers & Chemical Engineering, 19: 131–136

Westerlund T, Pettersson F, Grossmann I E (1994). Optimization of pump configurations as a minlp problem. Computers & Chemical Engineering, 18(9): 845–858

Wu C, Wang Y, Lu Z, Pardalos P M, Xu D, Zhang Z, Du D Z (2018). Solving the degree-concentrated fault-tolerant spanning subgraph problem by DC programming. Mathematical Programming, 169(1): 255–275

Zamora J M, Grossmann I E (1998a). A global minlp optimization algorithm for the synthesis of heat exchanger networks with no stream splits. Computers & Chemical Engineering, 22(3): 367–384

Zamora J M, Grossmann I E (1998b). Continuous global optimization of structured process systems models. Computers & Chemical Engineering, 22(12): 1749–1770

Zhang H, Wang S (2006). Global optimization of separable objective functions on convex polyhedra via piecewise-linear approximation. Journal of Computational and Applied Mathematics, 197(1): 212–217

Zheng Q P, Wang J, Pardalos P M, Guan Y (2013). A decomposition approach to the two-stage stochastic unit commitment problem. Annals of Operations Research, 210(1): 387–410

Zwick U (1999). Outward rotations: A tool for rounding solutions of semidefinite programming relaxations, with applications to max cut and other problems. In: Proceedings of the 31st Annual ACM Symposium on Theory of Computing, ACM, 679–687