On Solving Optimization Problems with Hidden Nonconvex Structures

Alexander S. Strekalovsky

1 Introduction

The most of real-life problems, according to Leibnitz and Euler [3, 10, 13], can be stated as optimization problems, because the lows of the nature just follow the principles of Fermat, Lagrange, Euler, and other equations provided by extremum principles.

On the other hand, the contemporary situation can be characterized by the crucial impact and the increasing value of the numerical methods in view of computational solving the problems of practical interest.

It is worthy to note that the optimization problems must be separated into two parts: convex and nonconvex. From the viewpoint of the numerical processing of the problem of a rather general kind

$$\begin{cases} f_0(x) \downarrow \min_x, & x \in S \subset \mathbb{R}^n, \\ f_i(x) \leq 0, & i = 1, 2, \ldots, m, \end{cases} \qquad (\mathcal{P}_0)$$

there exists a "solvable case"—this one of the convex optimization problems, those where the domain S and the functions f_0 and f_i are all convex [3, 13, 22, 40].

Under minimal additional computability assumptions a convex optimization problem is computationally tractable [3, 22]. It means that the computational effort required to solve the problem to a given accuracy grows moderately with the dimension of the problem and the required number of accuracy digits.

A.S. Strekalovsky (✉)
Institute for System Dynamics and Control Theory of SB RAS,
Lermontov St. 134, Irkutsk, 664033, Russia
e-mail: strekal@icc.ru

In contrast to this, general-type nonconvex problems are too difficult for numerical solution, since in a real-life nonconvex optimization problem there can exist a lot (often a huge!) of local extrema and stationary points which are rather far from a global solution [9, 17, 25, 28].

As a consequence, the classical optimization methods (conjugate gradients, Newton's and quasi-Newton's methods, TRM, SQP, IPM, etc.) turn out to be inoperative, in general, and ineffective as to finding a global solution in nonconvex problems because they are not able to escape a local pit.

Moreover, specialists in applied problems do not think about the correctness of direct application of classical optimization methods in nonconvex problems, while the numerical results are interpreted only in the content aspect, without thinking of the fact that all classical optimization methods converge to the global solution only in convex problems [3, 22].

At the same time, in nonconvex problems, the direct application of standard methods may have unpredictable consequences [3, 17, 24, 28, 40], and sometimes may even distract one from the desired solution. So, these arise various approaches which completely neglect the classical optimization methods and use a direct selection way employing, for example, the B&B idea or cut's method. As well known the latter algorithms suffer the curse of dimension, when the volume of computations grows exponentially side by side with the growth of the problem's dimension [17]. We are sure, there exists also another way of solving nonconvex problems of high dimension [18, 19, 26–37].

In the recent two decades, we have managed to construct a theory of global search, which is harmonic from the viewpoint of the theory of optimization and which unexpectedly has turned out to be rather efficient in the aspect of computations, especially for the problems of high dimensions. Simultaneously necessary and sufficient Global Optimality Conditions (GOCs) for the principal classes of nonconvex problems can be viewed as the kernel of the theory (see below) [28].

Furthermore, we have proposed a family of local search methods (LSMs), which, on the one hand, in some cases develop methods earlier known for the special problems and, on the other hand, this family of LSMs represents a joint ensemble of methods, which is harmonic from the viewpoint of GOCs [28, 31, 33].

Moreover, the procedures of escape from stationary or local solutions, which are based on GOCs, are unique and quite efficient even in case of any simplest implementation [18, 19, 26–28, 31–37].

Besides, the approach elaborated has been tested on a wide field of popular nonconvex problems (some part of which is represented below). It has demonstrated an unexpected efficiency during the numerical solving problems of high dimension. Note, convex optimization methods are successfully used "inside" the procedures of local and global search proposed [18, 19, 26–28, 31–37].

Finally, we have to add that, according to the opinion of numerous confirmed specialists in optimization, the most attractive and promising fields of investigation and, may be, even modelling paradigms in optimization in twenty-first century can be represented (see [24]), in particular, by the following examples which both possess the hidden nonconvex structures:

- the search for equilibriums in competitions (conflict situations or games);
- hierarchical optimization problems.

Unexpectedly for us, we turned out to be on this main stream, but rather prepared, i.e. possessing a suitable mathematical apparatus.

2 Examples of Applied Problems

2.1 Linear Complementarity Problems

As well known [7], the linear complementary problem (LCP) aims at finding the pair of vector $(x, w) \in \mathbb{R}^{n+n}$, which satisfy the following conditions:

$$\left. \begin{array}{cc} Mx + q = w, & \langle x, w \rangle = 0, \\ x \geq 0, & w \geq 0, \end{array} \right\} \tag{1}$$

for a given vector $q \in \mathbb{R}^n$ and a given real $(n \times n)$-matrix M which is, in general, indefinite. Many physical, engineering problems (the braking problem; the problem of contact; the problem of viscoelastic twisting, etc.), some economic problems (the problems of market equilibrium, the problem of optimal constant basic capital, etc.) and problems of computational geometry can often be stated as LSP. From the first glance any nonconvexity is not visible in (1). Even if we consider a similar formulation of LSP, for example,

$$\left. \begin{array}{cc} \langle x, Mx + q \rangle = 0, & M = M^\top, \\ x \geq 0, & Mx + q \geq 0, \end{array} \right\} \tag{1$'$}$$

a nonconvexity does not appear since all data stays to be linear. However, if one looks at the problem as optimization problem:

$$\left. \begin{array}{c} \Phi(x) := \langle x, Mx \rangle + \langle x, q \rangle \downarrow \min, \quad x \in S, \\ S \stackrel{\triangle}{=} \{ x \in \mathbb{R}^n \mid x \geq 0, \ Mx + q \geq 0 \}, \end{array} \right\} \tag{2}$$

it becomes clear that the properties and the structure of the LSP (1) depend on the features of the matrix M, as follows.

(a) If M is nonnegative definite, the problem (2) is convex, i.e. solvable with the classical methods, for instance, the conjugate gradient method (CGM).
(b) If M is negative definite, then the problem (2) turns out to be nonconvex (anticonvex) optimization problem of concave minimization (that is equivalent to convex maximization).
(c) If M is indefinite, i.e. it has positive and negative eigenvalues, then the problem (2) must be classified as a d.c. minimization problem:

$$\Phi(x) = g(x) - h(x) \downarrow \min_x, \quad x \in S, \tag{3}$$

where $g(x) = \langle x, M_1 x \rangle + \langle q, x \rangle$, $M_1 = M_1^\top > 0$, $h(x) = \langle x, M_2 x \rangle$, $M_2 = M_2^\top > 0$, $M = M_1 - M_2$, $g(\cdot)$ and $h(\cdot)$ are strongly convex functions on $\mathbb{R}^n$. In each of cases (a), (b), and (c) one has to apply the different methods of local and global search in order to find a global solution to Problem (2) or, what is equivalent, to find a solution to Problem (1).

The conclusion is unexpected: if anyone needs to have a solution to the LCP (1), then he has to choose the only one way (global search method (GSM)) among three different paths (GS methods) dependent on the properties of the matrix M. Below we will show how to do it.

So, very simple, from the first sight, Problem (1) may turn out to be very difficult to solve, since it possesses, in general, a hidden nonconvexity.

2.2 Search for an Equilibrium

As example of equilibrium problems, let us consider the bimatrix games [33] which reflect the conflict of two parties (players), each one having a finite number of strategies. After having introduced the mixed strategies, we obtain

$$
\left.
\begin{aligned}
\langle x, Ay \rangle &\uparrow \max_{x}, \quad x \in S_m, \\
\langle x, By \rangle &\uparrow \max_{y}, \quad y \in S_n, \\
S_p = \Big\{ x \in \mathbb{R}_+^p \mid \sum_{i=1}^{p} x_i = 1 \Big\}, \quad p &= m, n.
\end{aligned}
\right\}
\tag{4}
$$

Some economics, engineering and ecological problems can be represented in the form of bimatrix games, in which the Nash equilibrium is the common concept and can be represented as follows: find an equilibrium situation $(x^*, y^*) \in S_m \times S_n$:

$$
\left.
\begin{aligned}
\langle x^*, Ay* \rangle &\geq \langle x, Ay^* \rangle \quad x \in S_m, \\
\langle x^*, By^* \rangle &\geq \langle x^*, By \rangle \quad y \in S_n.
\end{aligned}
\right\}
\tag{5}
$$

In formulae (4) and (5), from the first sight, any nonconvexity also is not yet visible, since all the data is linear, and the problem (4)–(5) seems to be convex, i.e. solvable by the classical methods and approaches.

However, it turns out that the search for the Nash equilibrium can be reduced [20, 21] to solving the following nonconvex (in general) problem of mathematical programming:

$$
\left.
\begin{aligned}
F(x, y, \alpha, \beta) &:= \langle x, (A+B)y \rangle - \alpha - \beta \uparrow \max, \\
x^\top B - \beta e_n &\leq 0_n, \quad x \in S_m, \\
Ay - \alpha e_m &\leq 0_m, \quad y \in S_n,
\end{aligned}
\right\}
\tag{6}
$$

where $\alpha, \beta \in \mathbb{R}$, $e_p = (1, 1, \ldots, 1)^\top \in \mathbb{R}^p$, $p = m, n$. Note that the numbers α_* and β_* in a global solution $(x^*, y^*, \alpha_*, \beta_*)$ to Problem (6) are the optimal profits of the first

and second players, respectively, in the game (4)–(5), while the pair (x^*, y^*) turns out to be just a Nash equilibrium point in the game (4)–(5).

On account of the formulation (6) it becomes clear that a way (method) of finding a Nash point strongly depends on the properties of the matrix $(A + B)$.

So, the conclusion is obvious here and consists in the fact that the initial statement (4)–(5) of a bimatrix game is deceptive in the sense that it has, in general, a hidden (implicit) nonconvexity.

2.3 Hierarchical Optimization Problems

Hierarchical problems are encountered in practice because of impossibility of accumulation of the total available information at the upper level in the process of investigation of structurally complex control systems (social, economic, ecological-economic ones, etc.) and, as a consequence, possess some hidden non-convexity generated by just hierarchical structures.

For example, the financial systems in the economic power countries are usually constructed as bilevel systems. Besides, the electric energy system in USSR was organized as a four-level system.

As to mathematical aspects of the statement, problems of bilevel programming represent extremum problems, that side by side with standard constraints which are expressed in terms of equalities and inequalities, include the constraints described with the aid of optimization subproblem representing the lower level of the bilevel problem (or the player called the follower in difference with the player of the upper level called the leader).

To begin with, let us consider the linear bilevel problem

$$(\mathcal{LBP}): \quad \begin{cases} F(x,y) := \langle c, x \rangle + \langle d, y \rangle \downarrow \min_{x,y}, \\ x \in X = \{ x \in \mathbb{R}^m \mid Ax \leq b \}, \\ y \in Y_*(x) := \operatorname{Arg}\min_y \{ \langle d_1, y \rangle \mid y \in Y(x) \}, \\ Y(x) = \{ y \in \mathbb{R}^n \mid A_1 x + B_1 y \leq b_1 \}, \end{cases}$$

where $c \in \mathbb{R}^m$, $d, d_1 \in \mathbb{R}^n$, $b \in \mathbb{R}^p$, $b_1 \in \mathbb{R}^q$, and A, A_1, B_1 are matrices of corresponding dimensions. Suppose, that
$(\mathcal{H}_1)$: the function $F(x,y)$ is bounded below on the nonempty set Z,

$$Z := \{ x \in \mathbb{R}^m, \ y \in \mathbb{R}^n \mid Ax \leq b, \ A_1 x + B_1 y \leq b_1 \};$$

$(\mathcal{H}_2)$: the function $\langle d_1, y \rangle$ is bounded below on the set $Y(x)$ for all $x \in X$.
Even in this very simple case it is easy to construct an example showing the non-convexity of the problem $(\mathcal{LBP})$.

Example 1. ([8]) Consider the problem

$$F(x,y) = x + 3y \downarrow \min_{x,y}, \quad x,y \in \mathbb{R}, \left.\begin{array}{l} \\ \\ \end{array}\right\}$$
$$1 \leq x \leq 6, \quad y \in Y_*(x) = Sol(\mathcal{P}_L), \qquad (\mathcal{LBP}_1)$$

$$(\mathcal{P}_L): \quad \left\{\begin{array}{l} f(y) = -y \downarrow \min_y, \\ x + y \leq 8, \quad x + 4y \geq 8, \\ \quad x + 2y \leq 13. \end{array}\right.$$

Regardless of the convexity of the set

$$Z = \{(x,y) \in \mathbb{R}^2 \mid 1 \leq x \leq 6, \ x+y \leq 8, \ x+4y \geq 8, \ x+2y \leq 13\},$$

it is easy to see even geometrically that the set

$$Z_* = \{(x,y) \in Z \mid y \in Y_*(x)\}$$

is nonconvex which provides for the nonconvexity in the problem $(\mathcal{LBP}_1)$. □

So, the Example 1 shows the importance of the preliminary theoretical study of hierarchical optimization problems and, as, may be, the simplest case of $(\mathcal{BP})$, the $(\mathcal{LBP})$.

2.4 Problems of Financial and Medical Diagnostics

Such problems are well known as applied ones, and on the other hand, these problems are often interpreted as the problems of generalized separability. For example, if the two sets of points $\mathcal{A}$ and $\mathcal{B}$ are characterized by the matrices $A = [a^1, \ldots, a^M]$, $B = [b^1, \ldots, b^N]$, $a^i, b^j \in \mathbb{R}^n$, then the problem of polyhedral separability may be reduced to the problem of minimization of the nonconvex nondifferentiable error function ($V = (v^p)$, $\Gamma = (\gamma_p)$, $\gamma_p \in \mathbb{R}$, $v^p \in \mathbb{R}^n$, $p = 1, \ldots, P$)

$$F(V,\Gamma) = F_1(V,\Gamma) + F_2(V,\Gamma), \qquad (7)$$

$$F_1(V,\Gamma) = \frac{1}{M} \sum_{i=1}^{M} \max\{0; \max_{1 \leq p \leq P} (\langle a^i, v^p \rangle - \gamma_p + 1)\}, \left.\begin{array}{l} \\ \\ \\ \\ \end{array}\right\}$$
$$F_2(V,\Gamma) = \frac{1}{N} \sum_{j=1}^{N} \max\{0; \min_{1 \leq p \leq P} (-\langle b^i, v^p \rangle + \gamma_p + 1)\}. \qquad (8)$$

In this problem it is also not clear with which kind of nonconvexity we are dealing and how to overcome not only the nonsmoothness of the problem but also a nonconvexity generated apparently by $F_2(\cdot)$.

But, anyway, the question arises how to attack optimization problems with a hidden or an explicit nonconvexities.

3 Optimization Problems with the Functions of A.D.Alexandrov

The targets of our presentation can be bounded by consideration of the class $DC(\mathbb{R}^n)$ of the functions $f(\cdot)$ which can be represented as the difference of two convex functions (d.c. functions). The class was, for the first time, introduced in 1934 [1, 2] by the Russian mathematician A.D.Alexandrov, the member of AS of USSR.

Nowadays, this class is viewed by the specialists [9, 17, 25, 39] to be rather wise for consideration. Furthermore, the $DC(\mathbb{R}^n)$ possess several remarkable properties.

(a) The set $DC(\mathbb{R}^n)$ is generated by the well-studied class—the convex cone of convex functions and forms a linear space [12, 13, 17, 28, 39].
(b) $DC(\mathbb{R}^n)$ includes the well-known classes such as twice differentiable functions, power and trigonometric polynomials [12, 13, 17, 39].
(c) Any continuous function on a compact set $\mathcal{K} \subset \mathbb{R}^n$ can be approximated at any desired accuracy (in the topology of homogeneous convergence) by a function from $DC(\mathcal{K})$ [12]. Consequently, any optimization problem with continuous functions can be approximated at any desired accuracy by an extremum problem with functions of A.D.Alexandrov.

Only note, that, if $f(\cdot)$ is a d.c. function, then there exists an infinite number of d.c. representations of the form (3), for example, in the form of difference of strongly convex functions.

Closedness of the set $DC(\mathbb{R}^n)$ of functions of A.D.Alexandrov with respect to the majority of operations, which are used in optimization, is also essential from the optimization viewpoint. For example, a sum, a difference, the module, the maximum, the minimum, etc. of the family of d.c. functions occur also in the class $DC(\mathbb{R}^n)$.

Besides, the number of problems with d.c. functions is so large that the majority of the specialists, who have a long-time experience of solving problems of d.c. programming are sure [9, 12, 13, 17, 39] that all (or almost all) nonconvex optimization problems turn out to be really d.c. problems.

In this connection, the following statement of optimization problem can be viewed as rather general:

$$\left.\begin{array}{l} f_0(x) = g_0(x) - h_0(x) \downarrow \min_x, \ x \in S, \\ f_i(x) = g_i(x) - h_i(x) \leq 0, \ i = 1,\dots,m; \\ f_j(x) = g_j(x) - h_j(x) = 0, \ j = 1,\dots,N. \end{array}\right\} \tag{9}$$

Here g_i, g_j, h_i, h_j are convex functions and S is convex set from $\mathbb{R}^n$.

Apparently, almost all the specialists in optimization areas could estimate Problem (9) as very difficult and unsolvable by the existing approaches and methods even for the case of middle dimension (say, $n = 100,\dots,1{,}000$.)

Actually, even very simple (from the viewpoint of Problem (9)) the convex maximization quadratic problem on a box (which is a very particular case of (9)):

$$h(x) = \tfrac{1}{2}\langle x, Qx\rangle \uparrow \max_{x}, \quad Q = Q^T > 0, \left.\begin{array}{c}\\\\\end{array}\right\}$$
$$x \in S := \Pi = \{x \in \mathbb{R} \mid \alpha_i \le x_i \le \beta_i \ i = 1,\dots,n\} \tag{10}$$

is proved to be NP-hard [9]. Therefore, to begin with, let us simplify the situation and start with rather simple (from the first glance) nonconvex optimization problems.

1. *D.C. minimization*

$$(\mathcal{P}): \qquad\qquad f(x) = g(x) - h(x) \downarrow \min, \ x \in D, \tag{11}$$

where $g(\cdot)$, $h(\cdot)$ are convex functions, and D is a convex set, $D \subset \mathbb{R}^n$.

2. *D.C. constraint problem*

$$(\mathcal{DCC}): \qquad\qquad \begin{array}{c} f_0(x) \downarrow \min_{x}, \ x \in S, \\ F(x) = g(x) - h(x) \le 0, \end{array} \left.\begin{array}{c}\\\\\end{array}\right\} \tag{12}$$

where $g(\cdot)$ and $h(\cdot)$ are as above, $S \subset \mathbb{R}^n$, $f_0(\cdot)$ is a continuous function.

3. *Convex maximization*

$$h(x) \uparrow \max, \ x \in D, \tag{13}$$

(when $g \equiv 0$ in (11)).

4. *Reverse-convex constraint problem*

$$\begin{array}{c} f_0(x) \downarrow \min, \ x \in S, \\ h(x) \ge 0, \end{array} \left.\begin{array}{c}\\\\\end{array}\right\} \tag{14}$$

($g \equiv 0$ in (12)).

Note, that any quadratic optimization problem with arbitrary matrices occurs in the classification (11)–(14) or takes the form (9).

4 Global Search Methodology

Since in our approach the general global search procedure includes two principal parts:

(a) local search;
(b) procedures of escaping a critical point provided by a LSM; we are going, first, to consider special (for each class of d.c. programming problems) LSMs.

4.1 Local Search

The ideas of the most of LSM are rather simple and may consist in the consecutive solution of the (partially) linearized problems which for Problems (11)–(14) turn out to be convex. As a consequence, it becomes possible to apply classical convex optimization methods (Newtonians, CGM, TRM, etc.) in order to find a solution to linearized problems, i.e. within the framework of Local Search Schemes.

So, unlike that in well-known methods of the so-called Global Optimization (such as B&B, cuts methods), which, say, "deny and ignore" the modern and classical optimization methods, we insist on the obligatory, but "indirect" application of these methods.

For example, as regards the problem of d.c. minimization $(\mathcal{P})$–(11), the basic element, the "cornerstone" of the Global and LSMs is solving the following (linearized at a current iteration point $x^s \in D$) convex problem

$$(\mathcal{PL}_s): \qquad \Phi_s(x) := g(x) - \langle h'(x^s), x \rangle \downarrow \min_x, \quad x \in D, \qquad (15)$$

where $h'(x^s) = h'_s \in \partial h(x^s), \; s = 1, 2, \ldots$ is a subgradient of the convex function $h(\cdot)$ at the point x^s [13]. It is clear that in the differentiable case h'_s coincides with the usual gradient ∇h_s [13].

Furthermore, the LSM itself for $(\mathcal{P})$–(11) may consist in the consecutive solving (likewise in the method of "direct iterations") Problems $(\mathcal{PL}_s)$–(15). More precisely, given $x^s \in D$, we can find $x^{s+1} \in D$ as an approximate solution to $(\mathcal{PL}_s)$ by means of some suitable convex optimization method (for example, BFGS), or one of the packages of applied software (Xpress-MP, IBM CPLEX etc).

So, we produce the sequence $\{x^s\}$ according to the inequality:

$$\Phi_s(x^{s+1}) := g(x^{s+1}) - \langle h'(x^s), x^{s+1} \rangle \; \leq \; \inf_x \{g(x) - \langle h'(x^s), x \rangle \mid x \in D\} + \delta_s \quad (16)$$

where the sequence $\{\delta_s\}$ fulfils the condition

$$\sum_{s=0}^{\infty} \delta_s < +\infty, \; \delta_s > 0, \; s = 1, 2, \ldots.$$

It was rather surprising that the process in this case converges in the following sense.

Theorem 1. *Suppose the cost function of Problem $(\mathcal{P})$–(11) is bounded below, so that*
$$\mathcal{V}(\mathcal{P}) := \inf(f, D) \overset{\triangle}{=} \inf_x \{f(x) \mid x \in D\} > -\infty.$$

Then the sequence $\{x^s\} \in D$ generated by the rule (16) satisfies the following conditions.

(a) The number sequence $\{f_s\}, \; f_s = f(x^s)$ converges in the sense, as follows:

$$\lim_{s \to \infty} f_s = f_* \geq \mathcal{V}(\mathcal{P}). \qquad (17)$$

(b) $\lim_{s\to\infty}[\inf_x\{g(x) - g(x^{s+1}) + \langle h'(x^s), x^{s+1} - x^s\rangle \mid x \in D\}] = 0.$ (18)

or, what is the same (see $(\mathcal{PL}_s)$–(15)),

$$\lim_{s\to\infty}[\mathcal{V}(\mathcal{PL}_s) - \Phi_s(x^{s+1})] = 0,$$ (18′)

where

$$\mathcal{V}_s := \mathcal{V}(\mathcal{PL}_s) := \inf_x\{g(x) - \langle h'(x^s), x\rangle \mid x \in D\}$$ (19)

is the optimal value of the linearized problem $(\mathcal{PL}_s)$–(15)
  (c) If the function $h(\cdot)$ in (15) is strongly convex, then we have

$$\lim_{s\to\infty}\|x^s - x^{s+1}\| = 0.$$ (20)

  (d) Any limit point x_* of the sequence $\{x^s\}$ generated by LSM (16) is a solution of the following linearized problem

$$(\mathcal{PL}_*):\qquad \Phi_*(x) := g(x) - \langle y_*, x\rangle \downarrow \min_x,\quad x \in D,$$ (21)

where $y_* = h'(x_*) \in \partial h(x_*)$.

Note that very frequently for small dimension ($n \leq 7, 8, 10$) cases LSM (16) provides for a global solution to $(\mathcal{P})$–(11).

It is interesting that historically the particular case ($g \equiv 0$) of LSM (16) for differentiable convex maximization problem (13) has been proposed by Bulatov in 1969 [4] and can be represented in the modern form as follows:

$$\langle h'(x^s), x^{s+1}\rangle + \delta_s \geq \sup_x\{\langle h'(x^s), x\rangle \mid x \in D\}.$$ (22)

Besides, the well-known "power" method for finding the maximal eigenvalue of a symmetric positive definite matrix A, or what is the same, for solving the problem [38]

$$\langle x, Ax\rangle \uparrow \max_x,\quad \|x\| \leq 1,$$ (23)

turns out to be very particular case of LSM (22), when the linearized problem $(\mathcal{PL}_s)$ (with $g \equiv 0$) can be solved analytically. Note that the method (22) in Problem (23) converges to the global solution [38].

Thus, one can conclude that the idea of linearization with respect to the basic nonconvexity of a nonconvex problem has certainly some age. Anyway, it is worth noting to mention the works of the group of Pham Dinh Tao in which the idea of linearization with respect to the basic nonconvexity also demonstrated its effectiveness [14–16].

Furthermore, special methods of local search have been developed for the problems with d.c. constraints (12), (14) (see [31]). These methods have also been grounded, for example, on considering linearized problems of the form

$$\left.\begin{array}{l} g(x) - \langle h'(x^s), x \rangle \downarrow \min_x, \\ x \in S, \quad f_0(x) \leq \zeta_s = f_0(x^s), \end{array}\right\} \tag{24}$$

and the duality of Tuy [39].

4.2 Global Optimality Conditions

The second step in the global search methodology can be viewed as the most important one and even crucial, because the question is how to escape a critical point (provided by an LSM and that is not a global solution).

Such a procedure is substantiated by the theoretical basis produced with the help of the so-called GOC which for the case of d.c. minimization problem $(\mathcal{P})$–(11) takes the following form.

Theorem 2. *If z is a global solution to $(\mathcal{P})$, $z \in Sol(\mathcal{P})$, $\zeta := f(z)$, then*

$$(\mathcal{E}): \quad \begin{cases} \forall (y, \beta) \in \mathbb{R}^n \times \mathbb{R}: \quad h(y) = \beta - \zeta, \\ g(x) - \beta \geq \langle h'(y), x - y \rangle \ \forall x \in D. \end{cases} \tag{25}$$

Proof. Suppose, for some pair (y, β) satisfying (25) and a feasible point $\hat{x} \in D$ the inequality in (25) is violated

$$g(\hat{x}) < \beta + \langle h'(y), \hat{x} - y \rangle.$$

Then due to convexity of $h(\cdot)$ we have

$$f(\hat{x}) \overset{\triangle}{=} g(\hat{x}) - h(\hat{x}) < h(y) + \zeta - h(y) = f(z),$$

or $f(\hat{x}) < f(z)$. Thus, $\hat{x}$ is "better" than z, which contradicts to $z \in Sol(\mathcal{P})$. □

So, when selecting the "perturbation parameters" (y, β) satisfying (25) and solving the linearized problem (sf. (15))

$$\Phi_y(x) := g(x) - \langle h'(y), x \rangle \downarrow \min_x, \quad x \in D, \tag{26}$$

(where $y \in \mathbb{R}^n$ is not obligatory feasible!) we obtain a family of starting points $x(y, \beta)$ for a further (assume) local search.

Moreover, on each level $\zeta_k = f(z^k)$ it is not necessary to investigate all the pairs (y, β) satisfying (25), $\zeta_k = \beta - h(y)$, but it is sufficient to discover the violation of the variational inequality (25) only for one pair $(\hat{y}, \hat{\beta})$.

After that, one proceeds to the next iteration of the global search: $z^{k+1} := \hat{x}$, $\zeta_{k+1} := f(z^{k+1})$, and starts the procedure from the very beginning. So, the idea of the GSM becomes considerably more clear.

For the case of d.c. constraint problem (12) the character of GOC is a little bit different. It particular, the necessary conditions are rather far from the sufficient ones. More precisely, we have the result as follows:

Theorem 3. *Assume that in Problem* (*DCC*)–(*12*) *the following condition holds:*

$$(\mathcal{G}): \qquad \left.\begin{array}{l} \textit{there does not exist a solution } x_* \in S \\ \textit{to Problem (12) such that } F(x_*) < 0. \end{array}\right\} \qquad (27)$$

If a point $z \in S$ *is a global solution to Problem* (*12*) *such that* $F(z) = 0$, *then*

$$(\mathcal{E}_1): \qquad \left\{\begin{array}{l} \forall(y,\beta) \in \mathbb{R}^n \times \mathbb{R}: \quad \beta = h(y), \ \forall h'(y) \in \partial h(y) \\ g(x) - \beta \geq \langle h'(y), x - y \rangle \ \forall x \in S, \ f_0(x) \leq f_0(z). \end{array}\right. \qquad (28)$$

Proof. Suppose we find some parameters (y_0, β_0), $h'(y_0) \in \partial h(y_0)$ and a point $x_0 \in S$ such that

$$\beta_0 = h(y_0), \quad f_0(x_0) \leq f_0(z), \text{ and } g(x_0) - \beta_0 < \langle h'(y_0), x_0 - y_0 \rangle.$$

Then due to convexity of $h(\cdot)$ we obtain

$$0 < \beta_0 - g(x_0) + h(x_0) - h(y_0) = -F(x_0).$$

Hence, we have the feasible point $x_0 \in S$, $F(x_0) < 0 = F(z)$ with the property $f_0(x_0) \leq f_0(z)$. It means that x_0 is a solution to Problem (12) as well as the point z. The latter contradicts to the condition $(\mathcal{G})$–(27). □

A procedure of escaping a local pit can be conducted in a similar manner as it was explained after Theorem 2.

In the next subsection such a procedure will be precised for the case of d.c. minimization problem (11).

4.3 Global Search Methods

In order to deal with nonconvex optimization problems and, in addition, on the basis of the rather large computational experience [11, 18, 19, 26–28, 31–37] we propose three principles on which can be produced a search for a global solution to d.c. optimization problems of the forms considered above.

1. Linearization with respect to the basic nonconvexities of the problem under scrutiny and, consequently, the reduction of the original problem to a family of (partially) linearized problems.
2. Application of contemporary convex optimization methods for solving linearized problems and, as a consequence, "within" special LSMs.
3. Construction of "good" approximations (resolving sets) of the level surfaces/epigraph boundaries of convex functions.

Moreover, by working rather long time (about 30 years) on the field of nonconvex optimization, several practical rules have been elaborated, which can be represented as follows:

1. Never apply convex optimization methods directly.
2. Exact classification of the problem under scrutiny.
3. Application of special (for the class of problems to which belongs your problem) LSM, or (your problem's) specific methods.
4. Application of GSM specialized for the class which includes your problem.
5. Construction of suitable approximations of level surfaces (and the boundaries of the epigraphs) of convex functions with the aid of the experience obtained during solving similar problems.
6. Application of convex optimization methods for solving linearized problems and within the framework of special LSM.

These rules may be explained otherwise and by examples, following the instances.

1. Never apply CGM or BFGS if you are not convinced that your problem is convex.
2. Try to separate the data of your problem into two parts—convex and anticonvex.

For example, dealing with a quadratic function of the kind

$$q(x) = \frac{1}{2}\langle x, Qx \rangle,$$

where the matrix Q is indefinite, you have to separate the matrix $Q(n \times n)$ into a difference $Q = Q_1 - Q_2$ of two symmetric positive definite matrices $Q_i = Q_i^T > 0$, $i = 1, 2$. Note there exists infinity of such representations and several methods and ways to obtain its [28, 38].

Further, it is very important where your quadratic function $q(x)$ is situated—in the objective function or among the constraint's data, because depending on the situation you have different types of the problem to solve—d.c. minimization (11) or d.c. constraint problem (12), respectively. And as a consequence, you have to follow the different strategy (GSM, see below).

To demonstrate the effectiveness of these practical rules, let us consider the following example.

Example 2 (Incorrect Classification). Consider the problem

$$\left. \begin{array}{c} \varphi(x) = \sum\limits_{i=1}^{n} \ln(1 + x_i) \downarrow \min\limits_{x}, \\ x \in \Pi = \left\{ x \in \mathbb{R}^n \mid -\frac{1}{2} \le x_i \le 3 \right\} \subset \mathbb{R}^n. \end{array} \right\} \tag{29}$$

Obviously, the point $z = \left(-\frac{1}{2}, \ldots, -\frac{1}{2} \right)^T$ is the solution to the problem. Suppose, the current iterate is $x^k = (0, \ldots, 0)^T$

$$\nabla f(x) = \left(\frac{1}{1+x_1}, \ldots, \frac{1}{1+x_n} \right)^T, \quad \nabla f(x^k) = (1, \ldots, 1)^T,$$

$$\nabla^2 f(x) = \begin{bmatrix} -\frac{1}{(1+x_1)^2} & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & -\frac{1}{(1+x_n)^2} \end{bmatrix}, \quad \nabla^2 f(x^k) = \begin{bmatrix} -1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & -1 \end{bmatrix}.$$

The auxiliary problem of Newton's method

$$\Phi(d) = \sum_{i=1}^{n} d_i - \sum_{i=1}^{n} d_i^2 \downarrow \min, \ d \in \Pi = (\Pi - x^k)$$

has obviously the solution $d = (3, \ldots, 3)^T$ (take the case $n = 2$), which is a direction to the worst feasible point $x = (3, \ldots, 3)^T$. As a consequence, the iteration of the line-search method $x^{k+1} = x^k + t d^k$ cannot escape from $x^k = (0, \ldots, 0)^T$ and the process is stopped at x^k. Besides, note that the auxiliary problem conserves the nonconvex character of the original problem (29).

In contrast to the incorrectness of the above classification, let us look at the goal function of Problem (29) as a concave function, and correspondingly at the problem (29) as the concave minimization problem. Then we immediately conclude that we are dealing just with Problem (13) $(h(\cdot) = -\varphi(\cdot))$. Hence, we have to apply, first, the special LSM (16) where $g(x) \equiv 0$. So, beginning at arbitrary feasible point $x^0 \in \Pi$, we have to solve the linearized problem

$$(\mathcal{PL}_0): \quad \langle \nabla \varphi(x^0), x \rangle = -\langle \nabla h(x^0), x \rangle \downarrow \min_x, \ x \in \Pi \subset \mathbb{R}^n,$$

i.e.

$$\sum_{i=1}^{n} \frac{1}{1+x_i^0} \cdot x_i \downarrow \min_x \quad -\frac{1}{2} \le x_i \le 3, \ i = 1, \ldots, n$$

that provides for the global solution to the original problem (29)

$$x^1 = z \triangleq \left(-\frac{1}{2}, \ldots, -\frac{1}{2} \right)^T \in Sol(\mathcal{P}).$$

So, the special LSM (in one step!) has found the global solution to Problem (29).

□

Let us return now to the construction of a GSM (strategy) based on GOC presented in Theorem 2 and specialized only for Problem $(\mathcal{P})$–(11).

The basic stages of such a GSM (strategy) can be described as follows:

I. Find a critical point z by means of the special LSM ((16), for example).
II. Choose a number $\beta \in [\beta_-, \beta_+]$, where $\beta_- = \inf(g, D)$, $\beta_+ = \sup(g, D)$ can be approximated by rather rough estimates.
III. Construct an approximation

$$\mathcal{A}(\beta) = \{y^1,\ldots,y^N \mid h(y^i) = \beta - \zeta, \ i = 1,\ldots,N = N(\beta)\}$$

of the level surface of the function $h(\cdot)$.

IV. Beginning at every point y^i of the approximation $\mathcal{A}(\beta)$ find a feasible point u^i by means of the special local search algorithm (16).

V. Verify the VI (25) from GOC

$$g(u^i) - \beta \geq \langle h'(w^i), u^i - w^i \rangle \quad \forall i = 1,\ldots,N, \tag{30}$$

where w^i may be found as the projection of the point u^i onto the convex set

$$\mathcal{L}(h, \beta - \zeta) = \{x \in \mathbb{R}^n \mid h(x) \leq \beta - \zeta\}.$$

VI. If $\exists j \in \{1,\ldots,N\}$ such that (30) is violated, then set $x^{k+1} := u^j$ and return to Stage I. Otherwise change β and return to Stage III.

Example 3. Consider the problem

$$f(x) \downarrow \min, \quad x \in \mathbb{R}, \tag{31}$$

$$f(x) = \begin{cases} \frac{1}{4}x^4 - \frac{1}{2}x^2, & x \geq 0, \\ \frac{1}{2}x^4 - x^2, & x < 0. \end{cases} \tag{32}$$

The d.c. representation is here obvious, for example

$$f(x) = g(x) - h(x),$$

where

$$g(x) = \begin{cases} \frac{1}{4}x^4, & x \geq 0, \\ \frac{1}{2}x^4, & x < 0, \end{cases} \qquad h(x) = \begin{cases} \frac{1}{2}x^2, & x \geq 0, \\ x^2, & x < 0. \end{cases} \tag{33}$$

Let us choose the starting point $x_0 = 100$, while the global solution is $z = -1$, which can be readily seen.

(A) *Local search*

We have $s = 0$, $x_0 = 100$, $\nabla h(x_0) = x_0 = 100$, since

$$\nabla h(x) = \begin{cases} \nabla h_1(x) = x, & x \geq 0, \\ \nabla h_2(x) = 2x, & x < 0. \end{cases} \tag{34}$$

Then the linearized (convex) problem $(\mathcal{PL}_0)$–(15) takes the form

$$(\mathcal{PL}_0): \quad \Phi_0(x) = g(x) - \langle \nabla h(x_0), x \rangle = \frac{1}{4}x^4 - 100x \downarrow \min_x, \quad x \in \mathbb{R}.$$

To simplify the situation, in order to solve (PL_s) let us apply the Fermat rule instead of any numerical method. This yields

$$\nabla \Phi(x) = x^3 - 100 = 0.$$

Taking into account that $4^3 = 64$, $5^3 = 125$, let us risk to set $x_1 \approx 4.5$. Further, we have to solve the next linearized problem $(s = 1)$

$(\mathcal{PL}_1):$ $$\Phi_1(x) = \frac{1}{4}x^4 - 4.5x \downarrow \min_x, \quad x \in \mathbb{R},$$

and, as a consequence, the equation

$$\nabla \Phi_1(x) = x^3 - 4.5 = 0,$$

whence it follows that $x_2 \approx 1.7$.

For $s = 2$ consider the next linearized problem

$(\mathcal{PL}_2):$ $$\Phi_2(x) = \frac{1}{4}x^4 - 1.7x \downarrow \min_x, \quad x \in \mathbb{R},$$

and the corresponding equation

$$x^3 - 1,7 = 0,$$

that provides for the solution $x_3 \approx 1.2$. Hence, it is clear that $\{x^s\}$ tends to $z^0 = 1 \in \text{Arglocmin}$ (31).

(B) *Global search.*

Step 1. Thus, beginning at $x_0 = 100$ LSM provided for the point $z^0 = 1$, $\zeta_0 := f(z^0) = -\frac{1}{4}$.

Step 2. To begin with let us choose $\beta_0 = g(z^0) = g(1) = \frac{1}{4}$.

Step 3. Now we need to construct an approximation

$$\mathcal{A}_0 = \left\{ y^1, y^2, \dots, y^N \mid h(y^i) = \beta_0 - \zeta_0 = \frac{1}{4} - \left(-\frac{1}{4}\right) = \frac{1}{2} \right\}.$$

$$i = 1, \quad h_1(y) \overset{\triangle}{=} \frac{1}{2}y^2 = \frac{1}{2}, \quad y > 0, \quad y_1 = 1,$$

$$i = 2, \quad h_2(y) \overset{\triangle}{=} y^2 = \frac{1}{2}, \quad y < 0, \quad y_2 = -\frac{\sqrt{2}}{2}.$$

So, we obtain $\mathcal{A}_0 = \{ y_1 = 1, \; y_2 = -\frac{\sqrt{2}}{2} \}$.

Step 4. Further, we have to solve the linearized problems $(i = 1, 2)$

$(\mathcal{PL}_i):$ $$\Phi_i(x) = g(x) - \langle \nabla h(y_i), x \rangle \downarrow \min_x, \quad x \in \mathbb{R}.$$

a) $i = 1,$

$$\Phi_1(x) = \frac{1}{4}x^4 - \langle \nabla h_1(y_1), x \rangle = \frac{1}{4}x^4 - x \downarrow \min_x.$$

The Fermat rule provides for $u_1 = 1$.

$i = 2,$

$$\Phi_2(x) = \frac{1}{4}x^4 - \langle \nabla h_2(y_2), x \rangle = \frac{1}{4}x^4 - \langle 2y_2, x \rangle = \frac{1}{4}x^4 + \sqrt{2}x \downarrow \min_x, \quad x \in \mathbb{R},$$

whence it follows $\nabla \Phi_2(x) = x^3 + \sqrt{2} = 0$ that yields $u_2 = -(2)^{1/6}$.

b) $i = 1$,

$$\Phi_3(x) = \frac{1}{2}x^4 - \langle \nabla h_1(y_1), x \rangle = \frac{1}{2}x^4 - x \downarrow \min_x.$$

As above, one has $u_3 = \sqrt[3]{0.5}$.

Here, it is necessary to note that the points u_2 and u_3 are unacceptable because the initial data and the final results are incompatible, i.e.

a) $i = 2$, $g(x) = g_1(x) = \frac{1}{4}x^4$ when $x \geq 0$ meanwhile the solution $u_2 = -(2)^{1/6}$ is negative;

b) $i = 1$. $g(x) = g_2(x) = \frac{1}{2}x^2$ when $x < 0$, while $u_3 = \sqrt[3]{0.5} > 0$.

Further we consider the last case.

c) $i = 2$.

$$\Phi_4(x) = g_2(x) - \langle \nabla h_2(y_2), x \rangle = \frac{1}{2}x^4 - \langle 2y_2, x \rangle = \frac{1}{2}x^4 + \sqrt{2}x \downarrow \min_x.$$

It is easy to see that the Fermat rule $\nabla \Phi_4(x) = 2x^3 + \sqrt{2} = 0$ yields $u_4 = -\left(\frac{\sqrt{2}}{2} \right)^{1/3}$.

Now, we have to verify VI (30).

Step 5. a) $i = 1$.

$$g(u_1) - \beta_0 - \langle \nabla h(y_1), u_1 - y_1 \rangle = \frac{1}{4}u_1^4 = \frac{1}{4} - \langle y_1, u_1 - y_1 \rangle = \frac{1}{4} - \frac{1}{4} - 0 = 0.$$

b) $i = 2$.

$$g_2(u_2) - \beta_0 - \langle \nabla h_2(y_2), u_2 - y_2 \rangle = \frac{1}{2}u_4^4 - \frac{1}{4} - \langle 2y_2, u_4 - y_2 \rangle$$

$$= \frac{1}{2}\left(\frac{\sqrt{2}}{3} \right)^{4/3} - \frac{1}{4} + \langle \sqrt{2}, -\left(\frac{\sqrt{2}}{2} \right)^{1/3} + \frac{\sqrt{2}}{2} \rangle$$

$$= \frac{1}{2}\left[\left(\frac{1}{2} \right)^{2/3} - \frac{1}{2} \right] + \langle \sqrt{2}, \left(\frac{\sqrt{2}}{2} \right) - \left(\frac{\sqrt{2}}{2} \right)^{1/3} \rangle.$$

Without a computer it is rather difficult to decide about the sign of the latter expression. Suppose, our program was incorrect at this point, and we turned out to be unsuccessful to violate the VI (30). What do we have to do further? It is necessary to change β for another value, i.e. to loop to Step 2.

Step 2. Change β_0 for $\beta_1 = \frac{3}{4}$.

Step 3. We need a new set $\mathcal{A}_1$ of points y_i satisfying

$$h(y) = \left\{ \begin{array}{ll} \frac{1}{2}y^2, & y \geq 0, \\ y^2, & y < 0 \end{array} \right\} = \beta_1 - \zeta_0 = \frac{3}{4} - \left(-\frac{1}{4} \right) = 1,$$

whence it follows

$$i = 1, \quad y_1 = \sqrt{2},$$
$$i = 2, \quad y_2 = -1.$$

On account of (34) we have to solve the linearized (convex) problems $(i = 1, 2)$

$$(\mathcal{PL}_i): \qquad \Phi_i(x) = g(x) - \langle \nabla h(y_i), x \rangle \downarrow \min_x, \quad x \in \mathbb{R}.$$

Note that here it is sufficient to investigate only the case $(i = 1,$ a$)$ and $(i = 2,$ b$)$, because other two $(i = 1,$ b$)$ and $(i = 2,$ a$)$ turn out to be unacceptable, as above. $i = 1,$ a$)$

$$g_1(x) - \langle \nabla h_1(y_1), x \rangle = \frac{1}{4}x^4 - x\sqrt{2} \downarrow \min_x, \quad x \in \mathbb{R}.$$

With the help of Fermat rule one has

$$x^3 - \sqrt{2} = 0, \quad x \geq 0, \quad u_1 = (2)^{1/6}.$$

$i = 2,$ b$)$

$$g_2(x) - \langle \nabla h_2(y_2), x \rangle = \Phi_4(x) = \frac{1}{2}x^4 + 2x \downarrow \min, \quad x \in \mathbb{R},$$

which provides for

$$2x^3 + 2 = 0, \quad u_4 = -1.$$

Now, we need to verify VI (30). However, it is sufficient to consider only the case $(i = 2,$ b$)$ with $u_4 = -1, y_2 = -1$. Actually, in this case due to (34) we have

$$g_2(u_4) - \beta_1 - \langle \nabla h_2(y_2), u_4 - y_2 \rangle = \frac{1}{2}u_4^4 - 1 - \langle 2(-1), -1 - 1 \rangle = \frac{1}{2} - 1 - 0 = -\frac{1}{2} < 0.$$

The latter inequality means that GOCs have been violated, and, moreover, we were successful to "jump" out the local pit $z^0 = 1$ directly to the global solution $z^1 = -1$ by means of *Global Search Strategy (Method)*. $\square$

5 Numerical Solution of the Applied Problems

5.1 Linear Complementarity Problem

As it was said in Sect. 2, we have to look at the LCP (1) as the optimization problem (2). Besides, we will consider the most difficult case when (2) is nonconvex. More precisely, the matrix M in the statement (2) is indefinite, i.e. possesses positive and negative eigenvalues.

Note that the LCP (1) represents necessary optimality conditions for the problem

$$f(x) = \frac{1}{2}\langle x, Mx \rangle + \langle q, x \rangle \downarrow \min, \quad x \geq 0. \tag{35}$$

However, this problem cannot replace (2) because M is indefinite. As a consequence, $f(\cdot)$ in (35) can be unbounded below, while the objective function in (2) is nonnegative and takes the zero value only at a solution to Problem (2). It is clear that this provides an additional information for the computational process.

Now, let us describe the principal stage of the *Global Search Algorithm* (GSA).

0. *Classification*. Thus, we decide to classify LCP with an indefinite matrix M as a d.c. minimization problem (3) with the strongly convex functions $g(\cdot)$ and $h(\cdot)$:

$$\begin{aligned} g(x) &= \langle x, M_1 x \rangle + \langle q, x \rangle, \quad h(x) = \langle x, M_2 x \rangle, \\ M_i &= M_i^\top > 0, \quad i = 1, 2, \quad M = M_1 - M_2. \end{aligned}$$

I. The next stage is the local search. In order to do it, let us apply the LSM (16) which (for the LCP (2)) takes the form of the consecutive solutions of the following linearized problem

$$\left. \begin{aligned} \Phi_s(x) &= \langle M_1 x, x \rangle + \langle q - 2M_2 x^s, x \rangle \downarrow \min, \\ x &\geq 0, \quad Mx + q \geq 0. \end{aligned} \right\} \tag{36}$$

To the end of solving Problems (36) we used the well-known XPress solver, which was especially designed for solving convex quadratic and linear programming problems.

On the other hand, to organize rather effective testing of the LSM (16), the data of the $(n \times n)$ matrix M were randomly generated in the interval $[-n, n]$ (see [34]). So, the set of randomly generated LCPs of type (2) and of the dimension varying from $n = 2$ to $n = 1000$ has been formed. Moreover, for every LCP we used three different starting points. Further, the local solution process has been performed and analyzed on this field of test LCPs. In particular, the results enable us to observe the behavior of the method (15)–(16) and to choose appropriate starting points for global search ("good-bad" points, starting at which the LSM (15)–(16) does not provide for a global solution to Problem (2)).

Note separately that due to Theorem 1 the linearized problems (36) may be solved at a low accuracy at the first steps; further, the accuracy δ_s can be gradually improved ($\delta_s \downarrow 0$), for example, $\delta_0 = 0.1$, $\delta_{s+1} = 0.5\delta_s$ until the condition $\delta_s \leq \delta$ is fulfilled with a given accuracy $\delta > 0$. The results of computational testing of LSM (16) have been presented for the first time in [34] and after have been considerably improved (till the dimension $n = 1,000$ with a 3.4 GHz Pentium computer with 1 Gb of memory). The auxiliary linearized problems (36) have been solved by XPress solver.

On the basis of the analysis of the results of computational testing [34] one can conclude that the LSM (16) showed itself rather effective for LCP (2). Moreover, it was considerably more effective in comparison with X-Press solver, because the latter was unable to deal with nonconvex LCP (2) of dimension $n \geq 10$, meanwhile

the LSM (15)–(16) yielded a critical (feasible) point for (2) in all considered test problems till the dimension $n = 1,000$ with obligatory and considerable decreasing of the goal function $\Phi(\cdot)$ in (2) (see [34]).

So, LSM (15)–(16) can be applied in a GSA, although it is not able, in general, to reach a global solution.

Now we can pass to a global search described in Sect. 4. To begin with, first, one has to propose a numeric solution of the equation

$$h(y) \triangleq \langle y, M_2 y \rangle = \beta - \zeta_k, \ M_2 = M_2^T > 0,$$

where $\zeta_k := \Phi(z^k)$, $\beta \in [\beta_-, \beta_+]$, more precisely to construct an approximation

$$\mathcal{A}_k(\beta) = \{y^1, \ldots, y^N \mid h(y^i) = \beta - \zeta_k, \ i = 1, \ldots, N_k\}$$

of the level surface $U_k(\beta) = \{x \mid h(x) = \beta - \zeta_k\}$. The construction of such an approximation is a key point in the implementation of the global search.

Regardless of the importance of this procedure, the construction may be performed in rather simple fashion, for example,

$$y^i = \mu_i d^i, \quad i = 1, \ldots, N, \tag{37}$$

where d^i are elements of some set in $\mathbb{R}^n$, for instance, $d^i = e^i$, $\{e^1, \ldots, e^n\}$ being the Euclidian basis of $\mathbb{R}^n$, and the numbers μ_i are chosen as the roots of the quadratic equation $h(\mu_i d^i) = \beta - \zeta_k$ due to the quadratic structure of $h(\cdot)$.

In order to solve Problem (2) we used the approximations as follows:

$$\mathcal{R}_1 = \{y^i = \mu_i e^i, \ y^{i+n} = -y^i \mid i = 1, \ldots, n\},$$

$$\mathcal{R}_2 = \{y^i = z^k + \mu_i e^i, \ y^{i+n} = z^k - \mu_i e^i \mid i = 1, \ldots, n\},$$

where z^k is the current iteration point.

Besides, we also applied the third approximation using the form (37), where $d^i(i = 1, \ldots, n)$ have been produced as the solutions of the linear programs

$$\langle e^i, x \rangle \downarrow \min_x, \ x \geq 0, \ Mx + q \geq 0, \ i = 1, \ldots, n \tag{38}$$

and d^{n+1} is the solution of the similar problem

$$\langle e, x \rangle \downarrow \min_x, \ x \geq 0, \ Mx + q \geq 0, \tag{39}$$

with $e = (1, \ldots, 1)^T \in \mathbb{R}^n$.

The results of computational testing of the developed GSA have first been published in [34] and turned out to be rather promising for the test problems of dimension till 400.

Now we are having the software which is able to solve LCP (2) till the dimension 10^3 in 10–12 min and till the dimension 10^4 in 90–150 min, remember, by means of (only one) almost the same computer as it was used in [34], without applying any parallel technology.

In addition, in order to compare the efficiency of GSA with existing software the same series of randomly generated problems have been solved using the solver PATH [34], which was especially designed for solving the LCP (1). Note that all computational simulations have been carried out by students and postgraduate students.

So, the qualities of the programs implemented may vary significantly.

Nevertheless, we can conclude that the developed GSA proved to be rather effective for solving (nonconvex) LCPs with indefinite matrix M.

5.2 Bimatrix Games

Here we present the principal points of the numerical search for Nash equilibrium (NE) defined in (5) in the two-person game stated in (4). The computational algorithm has been developed on the foundation of the following result of Mills [21].

Theorem 4. *([21]) (i) A situation (x^*, y^*) is a Nash equilibrium in the bimatrix game $G(A,B)$ (4) if and only if (x^*, y^*) is a part of a global solution $(x^*, y^*, \alpha_*, \beta_*) \in \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^2$ to Problem (6).*

(ii) Moreover, α_* and β_* are the payoffs of the first and the second players, respectively:

$$\langle x^*, Ay^* \rangle = \alpha_*, \quad \langle x^*, By^* \rangle = \beta_*.$$

(iii) Finally, the optimization value of the goal function in Problem (6) is equal to zero

$$\mathcal{V}(6) = F(x^*, y^*, \alpha_*, \beta_*) = 0.$$

$\square$

5.2.1 Classification

In order to develop a numerical method for solving Problem (6) we have, first, to classify it as a nonconvex problem. Because all the constraints in (6) are linear, we have to decide about the features of the cost function of Problem (6). It can be readily seen that this function has the d.c. decomposition as follows:

$$F(x, y, \alpha, \beta) = h(x, y) - g(x, y, \alpha, \beta), \tag{40}$$

where

$$\left.\begin{array}{l} h(x, y) = \frac{1}{4}(\|x + Ay\|^2 + \|x + By\|^2) \\[2mm] g(x, y, \alpha, \beta) = \frac{1}{4}(\|x - Ay\|^2 + \|x - By\|^2) + \alpha + \beta \end{array}\right\} \tag{41}$$

are convex functions ($h(\cdot)$ on $\mathbb{R}^{m+n}$, $g(\cdot)$ on $\mathbb{R}^{m+n+2}$). In other words, we have the d.c. minimization problem $(\mathcal{P}_{\text{BM}})$ as follows:

$$\left.\begin{array}{c} F_0(x,y,\alpha,\beta) = -F(x,y,\alpha,\beta) = g(x,y,\alpha,\beta) - h(x,y) \downarrow \min_{x,y,\alpha,\beta}, \\ (x,\beta) \in X = \{x \in S_m, \quad \beta \in \mathbb{R} \mid x^\top B \le \beta e_n\}, \\ (y,\alpha) \in Y = \{y \in S_n, \quad \alpha \in \mathbb{R} \mid Ay \le \alpha e_m\}. \end{array}\right\} \quad (6')$$

$(\mathcal{P}_{\text{BM}})$:

It is easy to see that the cost function $F_0(x,y,\alpha,\beta)$ is nonnegative $(F_0(\cdot) \ge 0$, $F(\cdot) \le 0)$ on the feasible set of Problem $(\mathcal{P}_{\text{BM}})$–$(6')$.

In addition, if we denote

$$\alpha(y) := \max_{1 \le i \le m} (Ay)_i, \quad \beta(x) := \max_{1 \le j \le n} (x^\top B)_j, \quad (42)$$

then due to the necessity proof of Theorem 2 we can reformulate Theorem 2 in the contrapositive form as follows.

Theorem 5. *([28]) If a feasible tuple $(\hat{x}, \hat{y}, \hat{\alpha}, \hat{\beta})$ is not a global solution to Problem $(\mathcal{P}_{BM})$–$(6')$, then there exist some vectors $(u,v) \in S_m \times S_n$ and $(\bar{x}, \bar{y}) \in S_m \times S_n$, and a number γ such that*

$$\gamma - h(u,v) = \zeta := F_0(\hat{x}, \hat{y}, \hat{\alpha}, \hat{\beta}) > 0, \quad (43)$$

$$g(u,v,\alpha(v),\beta(u)) \le \gamma \le \sup(g,D), \quad (44)$$

$$g(\bar{x},\bar{y},\alpha(\bar{y}),\beta(\bar{x})) - \gamma < \langle \nabla_x h(u,v), \bar{x} - u \rangle + \langle \nabla_y h(u,v), \bar{y} - v \rangle. \quad (45)$$

$\square$

Applying just this result we will develop a GSM for finding a global solution to $(\mathcal{P}_{\text{BM}})$–$(6')$. The first step of this GSM is a local search algorithm which takes into account the bilinear structure of the cost function $\alpha + \beta - \langle x, (A+B)y \rangle$ of the Problem $(\mathcal{P}_{\text{BM}})$–$(6')$.

5.2.2 Local Search

First, let us repeat that LSMs play the important role in the processes of search for a global solution to nonconvex problems, since it provides for the so-called critical (stationary) points which may be considerably better than a simple feasible point. Moreover, if a starting point occurs rather closed to a global solution (as in the case of Newton method for solving systems of nonlinear equations), then an LSM is able to provide for the global solution.

Therefore, we have to pay our attention and considerable efforts to a creation (a design or a choice) and the substantiation of local search procedures.

For instance, for the case of Problem $(6')$ it might be possible to apply the LSM (15)–(16) taking into account the d.c. representation (40)–(41) and applying the corresponding methods of quadratic programming.

However, in this case the bilinear nature of Problem $(\mathcal{P}_{\text{BM}})$–$(6')$, the specific character of the cost function $\langle x, (A+B)y \rangle$, namely, its bilinearity, would be lost. Therefore, we propose to follow another way, more natural in the case, taking

into account the bilinear structure of the goal function $F_0(x,y,\alpha,\beta) = \alpha + \beta - \langle x, (A+B)y \rangle$. Combining the linearization idea and the separation of the variables into groups according to the statement of Problem $(\mathcal{P}_{BM})$–$(6')$, we obtain without alternative a procedure of consecutive solving the following two (linearized at a point $(u,v) \in \mathbb{R}^m \times \mathbb{R}^n$) problems

$$(\mathcal{PL}_x): \quad \begin{cases} \beta - \langle (A+B)v, x \rangle \downarrow \min_{(x,\beta)}, \\ (x,\beta) \in X = \{(x,\beta) \in S_m \times \mathbb{R} \mid x^\top B \le \beta e_n\}, \end{cases} \tag{46}$$

$$(\mathcal{PL}_y): \quad \begin{cases} \alpha - \langle u^\top(A+B), y \rangle \downarrow \min_{(y,\alpha)}, \\ (y,\alpha) \in Y = \{(y,\alpha) \in S_n \times \mathbb{R} \mid Ay \le \alpha e_m\}. \end{cases} \tag{47}$$

Unexpectedly enough, the procedure of consecutive solving the Problem $(\mathcal{PL}_x)$ and $(\mathcal{PL}_y)$ converges in the following sense.

Theorem 6. *([33]) The sequence of the tuples $(x^s, y^s, \alpha_s, \beta_s)$ generated by the LSM consisting in the consecutive fulfilling of the following inequalities*

$$\alpha_{s+1} - \langle x^s(A+B), y^{s+1} \rangle - \frac{\rho_s}{2} \le \inf_{(y,\alpha)} \{\alpha - \langle x^s(A+B), y \rangle \mid (y,\alpha) \in Y\}, \tag{48}$$

$$\beta_{s+1} - \langle x^{s+1}(A+B), y^{s+1} \rangle - \frac{\rho_s}{2} \le \inf_{(x,\beta)} \{\beta - \langle x, (A+B)y^{s+1} \rangle \mid (x,\beta) \in X\}, \tag{49}$$

converges to a quadruple $(\hat{x}, \hat{y}, \hat{\alpha}, \hat{\beta})$ satisfying the conditions as follows

$$\left. \begin{array}{ll} F_0(\hat{x}, \hat{y}, \hat{\alpha}, \hat{\beta}) \le F_0(\hat{x}, y, \alpha, \hat{\beta}) & \forall(y,\alpha) \in Y, \\ F_0(\hat{x}, \hat{y}, \hat{\alpha}, \hat{\beta}) \le F_0(x, \hat{y}, \hat{\alpha}, \beta) & \forall(x,\beta) \in X, \end{array} \right\} \tag{50}$$

provided that $\rho_s > 0$, $s = 0, 1, 2, \ldots$, $\sum\limits_{s=0}^{\infty} \rho_s < +\infty$. □

We will call, henceforth, such a point satisfying (50) a critical point of Problem $(\mathcal{P}_{BM})$–$(6')$. The LSM (46)–(49) has been tested on a rather large field of well-known test problems [33], and also on the various test problems especially constructed with the help of the idea from [5], by beginning the known games of small dimension (2×2, 3×3) and until the test-games of rather high size (say, $m = n = 1,000$).

Computational simulations certify unexpected effectiveness of the developed LSM that naturally depends on the method or a package of applied software (CPLEX) that was used for solving the linear problems (46), (47). Now we are able to perform LSM with the data $m = n = 10^6$ rather easily and effectively.

5.2.3 Global Search Algorithm

Recall that, in addition to local search, the basic stages of a GSM include an approximation of the level surface of the convex function $h(\cdot)$ (which creates the basic

nonconvexity in Problem $(\mathcal{P}_{BM})$–$(6')$), a solution of a linearized problem $(\mathcal{PL}_s)$, the verification of the VI (30) with $w^i \in U(h, \gamma - \zeta) = \{(x,y) \in \mathbb{R}^{m+n} \mid h(x,y) = \gamma - \zeta\}$, and finally a line search along the variable $\gamma \in \mathbb{R}$.

Taking into account the particularities of Problem $(6')$ the following modifications have been introduced into the general scheme of Global Search on the bases of Theorem 5.

1. Due to the properties of the cost function $F_0(x,y,\alpha,\beta)$ (see Theorem 4) the supplementary stopping criterion was introduced.
2. Two new parameters (q and v) have been introduced in order to control the speed and the accuracy of the algorithm [23].

Let us now describe the GSM for solving Problem $(\mathcal{P}_{BM})$–$(6')$ in a more algorithmic form.

Assume, we are given a starting feasible point $(x^0, y^0, \alpha_0, \beta_0) \in D = X \times Y$; number sequences $\{\tau_k\}$ and $\{\delta_k\}$, $k = 0, 1, 2, \ldots$, $\tau_k \downarrow 0$, $\delta_k \downarrow 0$ $(k \to \infty)$; a set of directions $Dir = \{(\bar{u}^1, \bar{v}^1), \ldots, (\bar{u}^N, \bar{v}^N) \in \mathbb{R}^{m+n}\}$ the bounds $\gamma_- \approx \inf(g, D)$, $\gamma_+ \approx \sup(g, D)$; and parameters $v \in]0, 1[$ and q.

Global Search Methods for $(\mathcal{P}_{BM})$–$(6')$.

Step 0. Set $k := 0$, $(\bar{x}^k, \bar{y}^k, \bar{\alpha}_k, \bar{\beta}_k) := (x^0, y^0, \alpha_0, \beta_0)$, $s := 0$, $p := 1$, $\gamma := \gamma_-$, $\Delta\gamma = (\gamma_+ - \gamma_-)/q$.

Step 1. Starting from $(\bar{x}^k, \bar{y}^k, \bar{\alpha}_k, \bar{\beta}_k) \in D$, move to τ_k—critical point $(x^k, y^k, \alpha^k, \beta^k) \in D$ by means of the special LSM (48)–(49). Set $\xi_k := F_0(x^k, y^k, \alpha_k, \beta_k) \leq F_0(\bar{x}^k, \bar{y}^k, \bar{\alpha}_k, \bar{\beta}_k)$.

Step 2. (*Stopping criterion*). If $\xi_k \leq \varepsilon$, where ε is the prescribed accuracy, then STOP: $(x^k, y^k) \in NE(G, \varepsilon)$.

Step 3. With the help of the point $(\bar{u}^p, \bar{v}^p) \in Dir$ $(p = 1, .., N)$ construct a point (u^p, v^p) such that $h(u^p, v^p) = \gamma - \xi_k$. Compute the numbers $\alpha_p := \max_{1 \leq i \leq m}(Av^p)_i$, $\beta_p := \max_{1 \leq j \leq n}(u^p B)_j$.

Step 4. If $g(u^p, v^p, \alpha_p, \beta_p) > \gamma + v\gamma$, then set $p := p + 1$ and return to Step 3. Else go to Step 5.

Step 5. Starting at the point $(u^p, v^p, \alpha_p, \beta_p)$ find a $2\tau_k$-critical point $(\hat{x}^p, \hat{y}^p, \hat{\alpha}_p, \hat{\beta}_p) \in D$ of Problem $(6')$ by means of special LSM.

Step 6. (*Stopping criterion*). If $F_0((\hat{x}^p, \hat{y}^p, \hat{\alpha}_p, \hat{\beta}_p) \leq \varepsilon$, then STOP. $(\hat{x}^p, \hat{y}^p) \in NE(G, \varepsilon)$.

Step 7. Find a δ_k-solution (x_0^p, y_0^p) to the level problem or, what is equivalent,

$$\langle \nabla_x h(x_0^p, y_0^p), \hat{x}^p - x_0^p \rangle + \langle \nabla_y h(x_0^p, y_0^p), \hat{y}^p - y_0^p \rangle + \delta_k$$
$$\geq \sup_{(x,y)} \{\langle \nabla_x h(x_0^p, y_0^p), \hat{x}^p - x \rangle + \langle \nabla_y h(x_0^p, y_0^p), \hat{y}^p - y \rangle \mid h(x,y) = \gamma - \zeta_k\}, \quad (51)$$

where $h(x_0^p, y_0^p) = \gamma - \zeta_k$.

Step 8. Compute

$$\eta_k(\gamma) = g(\hat{x}^p, \hat{y}^p, \hat{\alpha}_p, \hat{\beta}_p) - \gamma - \langle \nabla_x h(x_0^p, y_0^p), \hat{x}^p - \hat{x}_0^p \rangle - \langle \nabla_y h(x_0^p, y_0^p), \hat{y}^p - \hat{y}_0^p \rangle.$$

Step 9. If $\eta_k(\gamma) \geq 0$ and $p < N$, then set $p := p+1$ and loop to Step 3.

Step 10. If $\eta_k(\gamma) \geq 0$ and $p = N$, then set $\gamma := \gamma + \triangle \gamma$ and $p := 1$ and return to Step 3.

Step 11. If $\eta_k(\gamma) < 0$, then $k := k+1, (\bar{x}^{k+1}, \bar{y}^{k+1}, \bar{\alpha}_{k+1}, \bar{\beta}_{k+1}) := (\hat{x}^p, \hat{y}^p, \hat{\alpha}^p, \hat{\beta}^p)$, and return to Step 1.

Step 12. If $p = N$ and $\eta_k(\gamma) \geq 0 \forall \gamma \in [\gamma_-, \gamma_+]$ (i.e., the line search with respect to $\gamma \in [\gamma_-, \gamma_+]$ is finished), then stop.

The GSM presented above is not an algorithm, since some of its steps have not been described clearly, and must be precised. For instance, it is not clear how to find the pair (x_0^p, y_0^p) on Step 7, besides, how to construct a point (u^p, v^p) on the level surface of $h(\cdot)$: $h(u^p, v^p) = \gamma - \zeta_k$ with the help of a given direction $(\bar{u}^p, \bar{v}^p)$ on Step 3. As to the first problem of Step 7, it can be solved analytically for the quadratic function

$$h(x,y) = \frac{1}{4}(\|x + Ay\|^2 + \|x + By\|^2), \tag{52}$$

more precisely, the exact solution is given by the formula [32, 33]

$$(x_0^p, y_0^p) = t(\hat{x}^p, \hat{y}^p), \; t = \left[\frac{\gamma - \zeta_k}{h(\hat{x}^p, \hat{y}^p)} \right]^{\frac{1}{2}}.$$

The construction of point (u^p, v^p) satisfying $h(u^p, v^p) = \gamma - \zeta_k$ can be done in the similar way:

$$(u^p, v^p) = \lambda_p(\bar{u}^p, \bar{v}^p), p = 1, \ldots, N,$$
$$\lambda_p = \lambda_p(\zeta_k, \gamma) = \pm[\gamma - \zeta_k/h(\bar{u}^p, \bar{v}^p)]^{\frac{1}{2}}.$$

To the end of the solving Problem $(\mathcal{P}_{\mathrm{BM}})$–$(6')$ it was used the approximations of the level surface $U(h, \gamma - \zeta) = \{(x,y)| \; h(x,y) = \gamma - \zeta\}$ (see Step 3) constructed with the help of the following sets of directions

$$Dir1 = \{(e^i, e^j) \in \mathbb{R}^{m+n}| \; i = 1, \ldots, m, \; j = 1, \ldots, n\},$$

where $\{e^i\}$ is the Euclidian basis in $\mathbb{R}^m$ and $\{e^j\}$ is the basis in $\mathbb{R}^n$, respectively;

$$Dir2 = \{(e^i + x, e^j + y)| \; i = 1, \ldots, m, \; j = 1, \ldots, n\},$$

where (x,y) is a critical point provided by the special LSM;

$$Dir3 = \{(a^j + e_m, b^i + e_n)| \; i = 1, \ldots, m, \; j = 1, \ldots, n\},$$

where $a^j \in \mathbb{R}^n$ are the columns in A and $b^i \in \mathbb{R}^n$ are the rows in B, and $e_p = (1, \ldots, 1) \in \mathbb{R}^p, p = m, n$.

Note that the sets $Dir1, Dir2, Dir3$ have been selected as the most efficient ones after comparative computational experiments. But on the other hand, it is easy to

see that the number of points in the constructed approximations strongly depends on the size of the problem, i.e. is equal to $m \times n$.

So, the number of points in the approximations grows as q^2, where $q = \min\{m; n\}$.

It is clear that this moment makes it prohibited the numerical solution of Problem ($\mathcal{P}_{BM}$)–(6) of high dimension due to the excessive solution time.

In order to avoid this drawback it was employed some reducing procedure of the sets $Dir 1, 2, 3$ to sets with the number of points equal to $2(m+n)$ [23, 33].

5.2.4 Computational Simulations

The numerical experiments were conducted applying software programs implementing the GSAs described above. For all the problems, a starting point was chosen as follows:

$$x_i^0 = \frac{1}{m}, \ i = 1, 2, \ldots, m; \ y_j^0 = \frac{1}{n}, \ j = 1, 2, \ldots, n,$$
$$\alpha_0 = \max_i (Ay^0)_i, \ \beta_0 = \max_j (x^0 B)_j.$$

The computational simulations have been separated into several stages, and the first results of these experiments have been published in [23].

Further, the analysis of the results allowed us to conclude about some shortcomings of the software program developed. First of all, it was the solving method for linearized problems (46)–(47). Recall that to the end the simplex method program or the supporting cone method program was employed, which showed itself very excessive from the viewpoint of the solution time of problems of high dimension.

As a consequence, for solving the BM games of rather high dimension (up to $1,000 \times 1,000$) we decided to apply ILOG CPLEX 9.1(http://www-01.ibm.com/soft-ware/commerce/optimization/cplex-optimizer/index.html) especially oriented to LP problems. In addition, in order to create the worst conditions for the global search software the entries of matrices A and B have randomly been generated from the interval $[-n, n]$, where $n = m$.

The software programs of global search were run on Pentium 4, CPU 3 GHz with 512 Mb of RAM and have been implemented by post-graduated students without a long computational experience.

Nevertheless, the results of computational solving of BM games ($m = n$) can be viewed as rather promising from the point of view of analysis of numeric results of Table 1.

In Table 1, $m = n$ is the number of pure strategies of players 1 and 2, F_0 stands for the value of the goal function at the starting points, F_k is the corresponding value at the best obtained point, st is the number of iterations of GSAs (or, what is the same, the number of critical (stationary) points passed by GS algorithms), LP and Loc represent the number of linearized problems solved and the number of local search algorithm's applications, respectively.

Table 1 Computational results

$m = n$	F_0	F_k	st	LP	Loc	Time
200	40.3811	0	2	972	65	00:43.66
250	51.1617	0	4	4,843	321	5:36.30
300	60.1209	0	2	28	2	00:44.17
400	75.0987	0	6	16,168	978	59:05.49
500	75.3494	0	2	158	20	05:57.88
600	84.1025	0	2	82	7	07:33.92
700	89.0439	0	2	54	6	11:50.34
800	99.8335	0	2	48	3	15:28.70
900	100.0419	0	3	136	11	29:31.38
1,000	106.8368	0	3	178	18	45:04.34

It can readily be seen that in the cases of $m = n = 250$ and 400 it happened to randomly generate very difficult problems.

Despite these difficulties, all test-problems have successfully been solved that certifies on the computational effectiveness of the software program created on the basis of the GSA and the Global Search Theory.

Now, we are preparing to attack the bimatrix game of dimension $m = n = 10^4$ and the similar three-person-game of dimension $m = n = l = 5$, and 10.

5.3 Quadratic-Linear Bilevel Optimization

In this subsection we will consider the following problem of bilevel programming

$$(\mathcal{BP}): \qquad F(x,y) := \frac{1}{2}\langle x, Cx \rangle + \langle c, x \rangle + \frac{1}{2}\langle y, Cy \rangle + \langle c_1, y \rangle \downarrow \min_{x,y}, \qquad (53)$$

$$(x,y) \in X := \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^n \mid Ax + By \leq a, \quad x \geq 0\}, \qquad (54)$$

$$y \in Y_*(x) := \underset{y}{\operatorname{Arg\,min}}\{\langle d, y \rangle \mid y \in Y(x)\} \qquad (55)$$

$$Y(x) := \{y \in \mathbb{R}^n \mid A_1 x + B_1 y \leq b, \quad y \geq 0\}, \qquad (56)$$

where we are seeking an optimistic solution [6, 8], i.e. the upper level (x, leader) and the lower level (y, follower) are searching together (in cooperation) a common solution (x_*, y_*). Here, $c \in \mathbb{R}^m$, $d, c_1 \in \mathbb{R}^n$, $a \in \mathbb{R}^p$, $b \in \mathbb{R}^q$, and matrices C, C_1, A, B, A_1, B_1 are of corresponding dimensions. In addition, $C = C^\top > 0$, $C_1 = C_1^\top > 0$, so that the leader cost function is a convex quadratic function, while the follower goal function is linear.

Assume that

$(\mathcal{H})$: (i) the function $F(x,y)$ is bounded from below on X,

(ii) the function $\langle d,y \rangle$ is bounded from below on $Y(x) \ \forall x \in Pr(X)$.

It is clear, that, from the first glance, any nonconvexity is not visible in the formulation $(\mathcal{BP})$–(53)–(56). In order to put it explicit we apply the KKT-conditions for the follower problem (55)–(56):

$$d + vB_1 \geq 0, \quad v \geq 0, \quad A_1 x + B_1 y \leq b, \left.\begin{array}{c} \\ \end{array}\right\} \tag{57}$$
$$\langle d,y \rangle - \langle A_1 x - b, v \rangle = 0,$$

where v is the Lagrangian multipliers. Since the follower problem is also convex, the relations (57) are equivalent to the statement (55)–(56).

Let us replace, now, in $(\mathcal{BP})$ the follower problem by (57). It yields us the new problem

$$(\mathcal{P}): \qquad F(x,y) \downarrow \min_{x,y,v},$$
$$Ax + By \leq a, \quad A_1 x + B_1 y \leq b, \quad d + vB_1 \geq 0, \left.\begin{array}{c} \\ \\ \end{array}\right\} \tag{58}$$
$$\langle d,y \rangle = \langle A_1 x - b, v \rangle, \quad x \geq 0, \quad y \geq 0, \quad v \geq 0.$$

The following result establishes the relation between the $(\mathcal{BP})$ and Problem $(\mathcal{P})$.

Theorem 7. *([8]) For the pair (x^*, y^*) to be a global solution to Problem $(\mathcal{BP})$–(53)–(56), it is necessary and sufficient that there exists a vector $v^* \in \mathbb{R}^q$ such that the triple (x^*, y^*, v^*) is a global solution to Problem $(\mathcal{P})$–(58).* $\qquad \square$

Note, that, first, the relation exists only between the global solutions of Problems $(\mathcal{P})$ and $(\mathcal{BP})$, but it does not take place between local solutions or between local and global ones.

Second, it is easy to see that the feasible set of Problem $(\mathcal{P})$–(58) is nonconvex because of the presence of the bilinear equality-constraint in (58). Thus, Problem $(\mathcal{P})$–(58) turns out to be nonconvex.

Let us denote

$$H(x,y,v) := \langle d,y \rangle - \langle A_1 x - b, v \rangle \tag{59}$$

and introduce a μ-parametric family of problems as follows:

$$(\mathcal{P}(\mu)): \qquad F_1(x,y,v,\mu) = F(x,y) + \mu H(x,y,v) \downarrow \min_{x,y,v},$$
$$(x,y,v) \in D := \{(x,y,v) \mid Ax + By \leq a, \quad A_1 x + B_1 y \leq b, \left.\begin{array}{c} \\ \\ \end{array}\right\} \tag{60}$$
$$d + vB_1 \geq 0, \quad x \geq 0, \quad y \geq 0, v \geq 0\},$$

where $\mu > 0$ is a penalty parameter. If we rewrite the function $H(\cdot)$ in the form

$$H(x,y,v) = \langle d + vB_1, y \rangle - \langle A_1 x + B_1 y - b, v \rangle, \tag{59'}$$

then it becomes clear that

$$H(x,y,v) \geq 0 \quad \forall (x,y,v) \in D. \tag{61}$$

Furthermore, it can be readily seen that for a fixed value of μ Problem $(\mathcal{P}(\mu))$ is convex and quadratic with respect to the variables (x,y), and, besides, bilinear with respect to the variables x and v. So, Problem $(\mathcal{P}(\mu))$ can be called quadratic-bilinear, but anyway it stays to be nonconvex with the convex feasible set D (defined in (60)) and the nonconvex objective function $F_1(\cdot)$. Below we will show that $F_1(x,y,v,\mu)$ is a d.c. function.

Let us suppose $(x(\mu),y(\mu),v(\mu))$ be a solution to Problem $(\mathcal{P}(\mu))$–(60) for a given $\mu \in \mathbb{R}$. Further, denote $H[\mu] = H(x(\mu),y(\mu),v(\mu))$. Then the following relations between Problems $(\mathcal{P})$–(58) and $(\mathcal{P}(\mu))$–(60) take place.

(i) If the equality $H[\hat{\mu}] = 0$ holds for some value $\hat{\mu} \in \mathbb{R}$ and $(\hat{x},\hat{y},\hat{v}) = (\hat{x} = x(\hat{\mu}),\hat{y} = y(\hat{\mu}),\hat{v} = v(\hat{\mu}))$ is a solution to Problem $(\mathcal{P}(\hat{\mu}))$, then the triple $(\hat{x},\hat{y},\hat{v})$ is a solution to Problem $(\mathcal{P})$.

(ii) Moreover, for all $\mu > \hat{\mu}$ the equality $H[\mu] = H(x(\mu),y(\mu),v(\mu)) = 0$ holds, and, in addition, $(x(\mu),y(\mu),v(\mu))$ is a solution to Problem $(\mathcal{P})$.

In connection with these assertions we have results more suitable for computational uses.

Proposition 1. *([36]) Let $(x(\mu),y(\mu),v(\mu)) \in D$ be a τ_1-solution to Problem $(\mathcal{P}(\mu))$, and, besides,*

$$H(x(\mu),y(\mu),v(\mu)) \leq \tau_2.$$

Then

(i) $y(\mu)$ is a τ_2-solution to the follower problem (55)–(56) with parameter $x = x(\mu)$;

(ii) $(x(\mu),y(\mu))$ is an approximate τ_1-solution to Problem $(\mathcal{BP})$–(53)–(56). □

The above assertions allow us to apply the global search methodology developed in Sect. 4 for solving Problem $(\mathcal{P}(\mu))$–(60) and, as a consequence, for finding an approximate global solution to Problem $(\mathcal{BP})$–(53)–(56).

5.3.1 Local Search

It can be readily seen that Problem $(\mathcal{P}(\mu))$ can be rewritten in the following form

$$(\mathcal{P}(\mu)): \quad F_1(x,y,v) := \frac{1}{2}\langle x,Cx\rangle + \langle c,x\rangle + \frac{1}{2}\langle y,C_1y\rangle + \langle c_1,y\rangle$$

$$+ \mu\left[\langle d,y\rangle - \langle A_1x - b,v\rangle\right] \downarrow \min_{x,y,v}, \quad (62)$$

$$(x,y) \in Z := \{(x,y) \mid Ax + By \leq a, \quad A_1x + B_1y \leq b, \quad x \geq 0, \quad y \geq 0\}, \quad (63)$$

$$v \in V := \{v \mid d + vB_1 \geq 0, \quad v \geq 0\}. \quad (64)$$

On account of the assumptions $(\mathcal{H})$, it is easy to see that the cost function $F_1(\cdot)$ is bounded from below on the set $D = Z \times V$. Further, the statement (62)–(64) of Problem $(\mathcal{P}(\mu))$ suggests the idea of local search consisting in a consecutive solution of

the problem (62)–(64) with respect to the groups of variables; more precisely, in the case (62)–(64) , first, with respect to the pair (x, y) and, after that, with respect to the variables v, or in the inverse order.

Note that Problem $(\mathcal{P}(\mu))$ with a fixed value of the variable v becomes a convex quadratic optimization problem. On the other hand, for a fixed pair (x, y) we obtain a linear programming (LP) problem with respect to v. So, these auxiliary problems can be solved by standard software packages (CPLEX, X-Press, etc.)

Therefore, we can produce local search as it was done for the Bimatrix games.

Given some starting point $v_0 \in V$, we describe a so-called V-procedure as follows:

Step 0. Set $s := 0$, $v^s := v_0$.
Step 1. Find a $\frac{\rho_s}{2}$-solution (x^{s+1}, y^{s+1}) of the problem

$$\left. \begin{aligned} &\frac{1}{2}\langle x, Cx \rangle + \langle c, x \rangle + \frac{1}{2}\langle y, C_1 y \rangle + \langle c_1, y \rangle \\ &+ \mu\left[\langle d, y \rangle - \langle A_1 x - b, v^s \rangle\right] \downarrow \min_{x,y}, \quad (x,y) \in Z, \end{aligned} \right\} \qquad (\mathcal{PL}_s)$$

so that the following inequality holds

$$F_1(x^{s+1}, y^{s+1}, v^s) \leq \inf_{(x,y)}\{F_1(x, y, v^s) \mid (x,y) \in Z\} + \frac{\rho_2}{2}. \qquad (65)$$

Step 2. Find a $\frac{\rho_s}{2}$-solution v^{s+1} of LP problem

$$\langle d - A_1 x^{s+1}, v \rangle \downarrow \min_v, \quad v \in V, \qquad (\mathcal{LP}_s)$$

so that the following inequality is satisfied

$$F_1(x^{s+1}, y^{s+1}, v^{s+1}) \leq \inf_v\{F_1(x^{s+1}, y^{s+1}, v) \mid v \in V\} + \frac{\rho_2}{2}. \qquad (66)$$

Step 3. Set $s := s + 1$ and loop to Step 1.
Under the condition

$$\rho_s > 0, \quad s = 0, 1, 2, \ldots, \quad \sum_{s=0}^{\infty} \rho_s < +\infty,$$

we can prove, as it was in the Bimatrix games, that the numerical sequence $\{F_{1s} = F_1(x^s, y^s, v^s)\}$ generated by the V-procedure from above is converging.

Moreover, if $(x^s, y^s, v^s) \to (\hat{x}, \hat{y}, \hat{v})$, then the point $(\hat{x}, \hat{y}, \hat{v})$ turns out to be a critical point of Problem $(\mathcal{P}(\mu))$–(62)–(64) [35] or partially global solution to $(\mathcal{P}(\mu))$, i.e.

$$\left. \begin{aligned} F_1(\hat{x}, \hat{y}, \hat{v}) &\leq F_1(x, y, \hat{v}) \quad \forall (x,y) \in Z, \\ F_1(\hat{x}, \hat{y}, \hat{v}) &\leq F_1(\hat{x}, \hat{y}, v) \quad \forall v \in V. \end{aligned} \right\} \qquad (67)$$

Note that if a point $(\bar{x}, \bar{y}, \bar{v})$ is a local solution to Problem $(\mathcal{P}(\mu))$–(62)–(64), then $(\bar{x}, \bar{y}, \bar{v})$ turns out to be a critical point of Problem $(\mathcal{P}(\mu))$. Thus, the notion of critical point just introduced is really substantiated by and connected with the common notion of local solution. The similar to V-procedure so-called XY-procedure (starting at a point $(x_0, y_0) \in Z$) has also been studied and substantiated.

In order to test the developed LSM a rather large field of test-problems of the form $(\mathcal{BP})$–(53)–(56) has been constructed with the help of the idea of Calamai and Vicente [5], which provides for the bilevel problems with well-known properties, local and global solutions (even the numbers of which is known).

Now, a few words about the numerical testing of LSM.

First, the computational simulation was threefold:

(a) to choose a suitable value of the penalty parameter μ that provides for the equality $H(x(\mu), y(\mu), v(\mu)) = 0$ (the exact penalty [3, 22, 40]);
(b) to find starting points suitable for Global Search, i.e. from which the LSM was not able to reach a global solution;
(c) and finally, to compare two versions of LSM (with V- or XY-procedures).

Analyzing the testing results, one concluded that the computational time was rather short (less than 0.1 s), when the stopping criterion was satisfied at the accuracy $\tau = 10^{-4}$.

Furthermore, the value $\mu = 10$ of penalty parameter μ turned out to be sufficient to reach the equality $H(x(\mu), y(\mu), v(\mu)) = 0$ at a τ-critical point $(x(\mu), y(\mu), v(\mu))$. The targets (b) and (c) have been also reached.

Moreover, it should be specially noted the high rate of convergence of the XY- and V-procedures on the considered series of randomly generated problems, only two iterations were needed (starting from arbitrary feasible point) in order to get a critical point. So, the results of computational testing of the LSM were rather promising [35, 37].

5.3.2 Global Search

Let us repeat that the numerical test results showed that the special LSMs (V- and XY procedures) do not, in general, yield a global solution, even in problems of small sizes.

According to the methodology of Sect. 4, first we need to derive an explicit d.c. decomposition (if possible) of the cost function of the problem under scrutiny.

It is not hard to see that the goal function $F_1(x, y, v)$ of the problem $(\mathcal{P}(\mu))$ can be represented as a difference of two convex functions, for instance, as follows:

$$F_1(x, y, v) = g(x, y, v) - h(x, v), \tag{68}$$

where

$$
\left.
\begin{aligned}
g(x,y,v) &= \frac{1}{2}\langle x,Cx\rangle + \langle c,x\rangle + \frac{1}{2}\langle y,C_1y\rangle + \langle c_1,y\rangle \\
&\quad + \mu\left(\langle v,b\rangle + \langle y,d\rangle + \frac{1}{4}\|v - A_1x\|^2\right), \\
h(x,v) &= \frac{\mu}{4}\|v + A_1x\|^2
\end{aligned}
\right\}
\tag{69}
$$

are convex functions. Note that this d.c. decomposition is different with respect to these ones that was used in [35–37].

As it was noted above, the procedures of escaping critical points are based on GOC of Theorem 2 (see (25)) and employing the constructive (algorithmic) property of GOC. In the case of Problem $(\mathcal{P}(\mu))$ these GOCs take the form as follows:

$$
(x_*,y_*,v_*) \in Sol(\mathcal{P}(\mu)), \quad \zeta := F_1(x_*,y_*,v_*) \Longrightarrow
$$
$$
\forall (z,w,\gamma) \in \mathbb{R}^{m+q+1}: \quad h(z,w) = \gamma - \zeta, \tag{70}
$$
$$
g(x,y,v) - \gamma \geq \langle \nabla_{xv}h(z,w), (x,v) - (z,w)\rangle \quad \forall(x,y,v) \in D. \tag{71}
$$

Besides, if for some $(\hat{z},\hat{w},\hat{\gamma})$ in (70) and $(\hat{x},\hat{y},\hat{v}) \in D$

$$
g(\hat{x},\hat{y},\hat{v}) < \hat{\gamma} + \langle \nabla_{xv}h(\hat{z},\hat{w}), (\hat{x},\hat{v}) - (\hat{z},\hat{w})\rangle,
$$

i.e. the VI (71) is violated, then due to the convexity of $h(\cdot)$ it follows

$$
F_1(\hat{x},\hat{y},\hat{v}) < F_1(x_*,y_*,v_*).
$$

In other words, $(\hat{x},\hat{y},\hat{v}) \in D$ is "better" than (x_*,y_*,v_*).

Similarly to Sect. 4 and according to the methodology presented in Sect. 3 Problem $(\mathcal{P}(\mu))$ is decomposed into several simpler problems as follows:

(a) one-dimensional search along the variable γ;
(b) constructing the level surface approximation of the convex function $h(x,v)$, which does not depend on y, as it was in our earlier papers [35–37]. It is clear that in this case the approximations must be easier to construct.

On the other hand, we have to pay attention to the fact that in view of the different d.c. representation (68)–(69) the global search has to be changed and becomes different with respect to [35, 36].

Assume, we are given a point $(x_0,y_0,v_0) \in \mathbb{R}^{m+n+q}$, numerical sequences $\{\tau_k\}$, $\{\delta_k\}$, $\tau_k, \delta_k > 0$, $k = 0,1,\ldots$, $\tau_k \downarrow 0$, $\delta_k \downarrow 0$ $(k \to \infty)$, numbers $\gamma_- \approx \inf_{(x,y,v)}(g,D)$ and $\gamma_+ \approx \sup_{(x,y,v)}(g,D)$, an algorithm's parameter M and a direction's set of the form

$$
Dir = \left\{(a^l,c^l) \in \mathbb{R}^{m+q} \mid (a^l,c^l) \neq 0, \ l = 1,\ldots,N\right\}.
$$

The GS algorithm used here can be represented as follows:

Step 0. Set $k := 0$, $(\bar{x}^k,\bar{y}^k,\bar{v}^k) := (x_0,y_0,v_0)$, $l := 1$. $\gamma := \gamma_-$; $\Delta\gamma := \gamma_+ - \gamma_-/M$.
Step 1. Starting at the point $(\bar{x}^k,\bar{y}^k,\bar{v}^k)$ construct a τ_k-critical point $(x_k,y_k,v_k) \in D$ in Problem $(\mathcal{P}(\mu))$ by applying V- or XY-procedure. Set $\zeta_k := F_1(x_k,y_k,v_k)$.

Step 2. Given a point $(a^l, c^l) \in Dir$, construct a point (z^l, w^l) such that $h(z^l, w^l) = \gamma - \zeta_k$.

Step 3. Solve the linearized problem as follows:

$$(\mathcal{PL}_l): \quad g(x, y, v) - \langle \nabla_{xv} h(z^l, w^l), (x, v) \rangle \downarrow \min_{(x,y,v)}, \quad (x, y, v) \in D.$$

Let the point $(\hat{x}, \hat{y}, \hat{v})$ be a solution to $(\mathcal{PL}_l)$.

Step 4. Starting at the point $(\hat{x}, \hat{y}, \hat{v})$ construct a δ_k-critical point $(\hat{x}_l, \hat{y}_l, \hat{v}_l)$.

Step 5. If $F_1(\hat{x}_l, \hat{y}_l, \hat{v}_l) < \zeta_k \overset{\triangle}{=} F_1(x_k, y_k, v_k)$, then set $(\bar{x}^{k+1}, \bar{y}^{k+1}, \bar{v}^{k+1}) := (\hat{x}_l, \hat{y}_l, \hat{v}_l)$, $k := k+1$, $l := 1$, $\gamma := \gamma_-$ and loop to Step 1.

Step 6. If $F_1(\hat{x}_l, \hat{y}_l, \hat{v}_l) \geq \zeta_k$ and $l < N$, then set $l := l+1$ and return to Step 2.

Step 7. If $F_1(\hat{x}_l, \hat{y}_l, \hat{v}_l) \geq \zeta_k$ and $l = N$, then set $\gamma := \gamma + \Delta\gamma, l := 1$ and come back to Step 2.

Step 8. If $l = N$, $F_1(\hat{x}_l, \hat{y}_l, \hat{v}_l) \geq \zeta_k \ \forall \gamma \in [\gamma_-, \gamma_+]$ (i.e., one-dimensional search along γ over the interval $[\gamma_-, \gamma_+]$ is terminated), then STOP; (x_k, y_k, v_k) is a critical point provided by Algorithm of global search.

Remark 1. It is clear that different values of the parameter M are responsible for the partitioning of the interval $[\gamma_-, \gamma_+]$ into a suitable number of parts to implement a passive one-dimensional search along γ. On the other hand, it is necessary to precise how to construct a direction's set Dir and, furthermore, an approximation of the level surface $h(z, w) = \gamma - \zeta_k$.

Taking into account that in contrast to the earlier papers [35–37] here due to (69)

$$h(x, v) \overset{\triangle}{=} \frac{\mu}{4} \|v + A_1 x\|^2, \tag{69'}$$

we have to choose $\gamma \geq \zeta_k$ so that γ_- can always be chosen as follows: $\gamma_- := \zeta_k$. Other points of the implementation of the algorithm were similar to [35–37].

Let us focus now on the construction of approximation of the level surface

$$U(\gamma) = \{(z, w) | h(z, w) = \gamma - \zeta_k\}.$$

Recall that, on the one hand, such an approximation should be representative enough to escape a critical point (if possible). On the other hand, if we are rather far from a global solution, then the approximation must allow us to "jump out" the critical point where we are.

Let us show how to construct an approximation. Given a set of directions

$$Dir = \{(a^l, c^l) \in \mathbb{R}^{m+q} | (a^l, c^l) \neq 0, l = 1, \ldots, N\},$$

we construct a point of an approximation $\mathcal{A}_n$ in a rather simple manner as follows:

$$(z^l, w^l) = \lambda_l(a^l, c^l), \ h(z^l, w^l) = \gamma - \zeta_k, l = 1, \ldots, N. \tag{72}$$

Due to (69 ′) the corresponding equation

$$\frac{\mu}{4}\|c^l + A_1 a^l\|\lambda_l^2 = \gamma - \zeta_k \qquad (72\,′)$$

leads us to very simple computing in order to calculate λ_l.

As the sets of directions, one can consider, for example, the set

$$Dir1 = \{(x^k + e^i, v^k + e^j), (x^k - e^i, v^k - e^j) \mid i = 1, \dots, m, j = 1, \dots, q\}$$

where (x^k, v^k) is the part of the current critical point (x^k, y^k, v^k); $e^i \in \mathbb{R}^m$, $e^j \in \mathbb{R}^q$ are the Euclidean basis vectors. Further, it has been employed the set

$$Dir0 = \{(a^i, b^j) \mid (a^i, b^j) \neq 0, \ i = 1, \dots, m, j = 1, \dots, q\}.$$

where a^i and b^j are, respectively, rows and columns of the matrix A_1, which specifies nonconvexity in the goal function of Problem $(\mathcal{P}(\mu))$–(60).

Note that the numbers of points in the approximations constructed are equal to $2qm$ and grow rapidly with the dimension. Therefore, we have also made the reduction of the approximations as described in [35, 36]. The first stages of computational testings of the developed GSA have been presented in [36, 37].

Here, we will show the preliminary results of further computational experiments with improved GSA described above (see Tables 2 and 3).

In particular we see in Table 2 the comparison of the results of computational solving the test-problems (generated, as above, with the help of the methodology from [5]) by GSA described above and by means of very popular package of applied software KNITRO (www.ziena.com/knitro.htm).

Since KNITRO is not able to solve bilevel problems directly, the equivalent formulation $(\mathcal{P}(\mu))$ (with $\mu = 10, 15, 20$) was used to this end.

On the other hand, GSA has been run on a computer with the processor Intel Core 2 Duo 2.0 GHz, while KNITRO used a computer with more powerful processor Intel Core 2 Quad 2.8 GHz. In Table 2 F_* is the known optimal values of the test-problems, F_{Kms} and T are the best values of the goal function and the corresponding solving time provided by KNITRO, while F_{XY}, F_V, and T stand for the best values of the cost function and the solution time obtained by GSA (using XY- or V-procedure as LSM). The bold values in Table 2 denote the successful cases when the known global solutions to the test-problems have been reached by the used algorithms.

Analyzing results of Table 2, it is easy to note that the KNITRO (multistart) was successful to find the global solutions only in 61 % of the test-problems of the middle dimension with the accuracy $\varepsilon = 10^{-2}$. Meanwhile, applying the programs implementing GSA, all considered test-problems have been solved at the same precision.

Moreover, it is not hard to see the big difference in solution time between GSA and KNITRO for the problems of middle dimension more than 10. For example, for $m = n = 30$, KNITRO worked about 1.5–2 h without reaching a global solution, meanwhile GSA provided for a global solution in 2 min approximately.

Table 2 Comparison of global search algorithm (GSA) with KNITRO

Name	F_*	KNITRO multistart		Global search			
		F_{Kms}	T	F_{XY}	T	F_V	T
5×5_1	−21	**−21**	7.7	**−21**	11.4	**−21**	7.3
5×5_2	−9	**−9**	8.7	**−9**	11.3	**−9**	4.7
5×5_3	−5	**−5**	9.9	**−5**	11.7	**−5**	4.6
10×10_1	−38	−30	1:17.2	**−38**	33.1	**−38**	23.7
10×10_2	−26	**−26**	1:17.0	**−26**	29.2	**−26**	12.4
10×10_3	−14	**−14**	1:22.4	**−14**	24.6	**−14**	16.0
15×15_1	−19	**−19**	11:08.5	**−19**	20.3	**−19**	19.0
15×15_2	−27	−19	5:43.6	**−27**	38.6	**−27**	31.5
15×15_3	−43	−35	7:01.2	**−43**	48.1	**−43**	35.5
20×20_1	−24	**−24**	19:13.7	**−24**	26.9	**−24**	45.9
20×20_2	−48	**−48**	30:20.5	**−47.999**	1:12.0	**−47.999**	1:10.9
20×20_3	−52	−32	29:54.4	**−52**	1:13.1	**−52**	52.0
30×30_1	−142	−134	1:34:34.9	**−141.997**	4:30.3	**−141.997**	1:46.8
30×30_2	−58	−38	1:31:23.5	**−58**	1:26.2	**−58**	1:51.3
30×30_3	−42	−29.999	2:39:15.4	**−42**	51.5	**−42**	1:08.9

Table 3 Testing of global search algorithm (GSA) on problems of high dimension

$m+n$	N	$LocSol_{avg}$	Loc_{avg}	St_{avg}	T_{avg}
20	1,000	146.2	2,012.7	1.9	8.57
40	1,000	13, 1284.1	3,436.6	2.1	20.69
60	100	$1.34 \cdot 10^8$	4,601.5	2.0	34.36
80	100	$1.17 \cdot 10^{11}$	6,485.1	2.1	59.51
100	100	$1.20 \cdot 10^{14}$	9,352.5	2.1	1:40.29
150	10	$3.78 \cdot 10^{21}$	8,050.3	3.0	1:52.38
200	10	$1.27 \cdot 10^{29}$	12,263.8	2.8	4:00.69
250	10	$4.27 \cdot 10^{36}$	17,704.3	2.7	8:04.28
300	10	$3.93 \cdot 10^{44}$	72,245.6	17.9	49:53.48
350	10	$2.12 \cdot 10^{52}$	216,721.1	25.2	3:56:09.12
400	10	$8.64 \cdot 10^{59}$	318,448.7	27.6	9:10:24.43

Now, let us look at Table 3 where presented the results of computational solution of the test-problems of high dimension (until $m = n = 200$) provided by the software program implemented in a computer with the processor Intel Core i5-2400 3.1 GHz. In Table 3 N is the number of test-problems in series, $LocSol_{avg}$ is an average number of local solutions which are not global in one problem of the series (this is very important difficulty index of the problem); Loc_{avg} stands for the average number of switching on of the LSM in conducting GSA; St_{avg} is the average number of iterations of GSA or critical points passed by GSA; T_{avg} is the average working time of the program implementing the GSA.

From the results of computation testing we can see, firstly, that all 2,350 randomly generated problems have been successfully solved so that, regardless the fantastic difficulty of the test-problems ($m + n = 300$ and $LocSol_{\text{avg}} = 3.93 \cdot 10^{44}$, $m + n = 400$ and $LocSol_{\text{avg}} = 8.64 \cdot 10^{59}$), GSA has found a global solution in every considered test-problem.

However, this version of program is characterized by the rapid increase in computing time with the growth of the dimension: for $m + n = 400$ it takes more than 9 h. On the other hand, it can be explained by the number Loc_{avg} of local search applications which is more than 310 thousands.

Moreover, it is not hard to note that the number St_{avg} of approximately critical points, at which it happened an improvement of the cost function, turned out to be rather moderate with respect to the number $LocSol_{\text{avg}}$ of local solutions (different from global ones) which is varying from rather big ($m + n = 60$ and $LocSol_{\text{avg}} = 1.34 \cdot 10^8$) until incalculable ($m + n = 300$, $LocSol_{\text{avg}} = 3.93 \cdot 10^{44}$; $m + n = 400$, $LocSol_{\text{avg}} = 8.64 \cdot 10^{59}$).

So, we conclude that the new results of computational solving the bilevel problem can be viewed as rather promising and competitive. Moreover, we did not be successful to find, at present, the solution's results of similar problems of such dimensions in the existing literature.

6 Concluding Remarks

In the present paper, new procedures of finding the solution to the linear complementarity problem with indefinite matrices, the Nash equilibrium in bimatrix games, and optimistic solution in quadratic-linear bilevel optimization problems have been proposed, discussed, and illustrated.

Further, a new approach based on GOCs, LSMs and GSMs, was applied in order to solve all three problems. In addition, the new results of computational solutions were presented in the paper. According to these results, the new approach has shown itself rather promising and competitive.

References

1. Alexandrov, A.D.: On surfaces represented a difference of convex functions. Proc. Acad. Sci. KSSR. Ser. Math. Mech. **3**, 3–20 (1949) (in Russian)
2. Alexandrov, A.D.: The surfaces that can be represented by a difference of convex functions. Proc. Acad. Sci. USSR. Ser. Math. Mech. **72**(4), 613–616 (1950)
3. Bonnans, J.F., Gilbert, J.C., Lemaréchal, C., Sagastizábal, C.A.: Numerical Optimization: Theoretical and Practical Aspects, 2nd edn. Springer, Berlin, Heidelberg (2006)
4. Bulatov, V.P.: The approximation method for solving some problems of mathematical programming. Applied Mathematics. Irkutsk State University, pp. 82–88 (1969)
5. Calamai, P.H., Vicente, L.N.: Generating quadratic bilevel programming test problems. ACM Trans. Math. Softw. **20**, 103–119 (1994)

6. Colson, B., Marcotte, P., Savard, C.: An overview of bilevel optimization. Ann. Oper. Res. **153**(1), 235–256 (2007)
7. Cottle, R.W., Pang, J.-S., Stone, R.E.: The linear complementarity problem. SIAM, Philadelphia (2009) [Originally published by Academic Press, Boston (1992)]
8. Dempe, S.: Foundations of Bilevel Programming. Kluwer, Dordrecht (2002)
9. Floudas, C.A., Pardalos, P.M. (eds.): Frontiers in Global Optimization. Kluwer, New York (2004)
10. Grötschel, M. (ed.): Optimization Stories. Documenta Mathematica, Bielefeld (2012)
11. Gruzdeva, T.V., Strekalovsky, A.S., Orlov, A.V., Druzhinina, O.V.: Nonsmooth minimization problems for the difference of two convex functions. Numer. Methods Program. **12**(2), 139–151 (2011) (in Russian)
12. Hiriart-Urruty, J.-B.: Generalized Differentiability, Duality and optimization for Problems dealing with Diffrence of Convex Functions. Lecture Notes in Economics and Mathematical Systems, vol. 256, pp. 37–69. Springer, Berlin (1985)
13. Hiriart-Urruty, J.B., Lemaréchal, C.: Convex Analysis and Minimization Algorithms. Springer, Berlin (1993)
14. Hoai An, L.T., Tao, P.D.: The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems. Ann. Oper. Res. **133**, 23–46 (2005)
15. Hoai An, L.T., Tao, P.D.: On solving linear complemetarity problems by DC programming and DCA. J. Comput. Optim. Theory Appl. **50**(3), 507–524 (2011)
16. Hoai An, L.T., Tao, P.D., Van Thoai, N., Canh, N.N.: DC programming techniques for solving a class of nonlinear bilevel programs. J. Glob. Optim. **44**(3), 313–337 (2009)
17. Horst R., Tuy H.: Global Optimization: Deterministic Approaches. Springer, Berlin (1993)
18. Malyshev, A.V., Strekalovsky, A.S.: Connection of some bilevel and nonlinear optimization problems. Russian Math. **55**(4), 83–86 (2011)
19. Malyshev, A.V., Strekalovsky, A.S.: On global search for pessimistic solution in bilevel problems. Int. J. Biomed. Softw Comput. Hum. Sci. (Special Issue on Variational Inequality and Combinatorial Problems) **18**(1), 57–61 (2011)
20. Mangasarian, O.L.: Equilibrium points in bimatrix games. J. Soc. Ind. Appl. Math. **12**, 778–780 (1964)
21. Mills, H.: Equilibrium points in finite games. J. Soc. Ind. Appl. Math. **8**(4), 397–402 (1960)
22. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, Berlin (2006)
23. Orlov, A.V., Strekalovsky A.S.: Numerical search for equilibria in bimatrix games. Comput. Math. Math. Phys. **45**(6), 947–960 (2005)
24. Pang, J.-S.: Three modelling paradigms in mathematical programming. Math. Program Ser B **125**(2), 297–323 (2010)
25. Pardalos, P.M., Resende, M.G.C. (eds.): Handbook of Applied Optimization. Oxford Univesity Press, New York (2002)
26. Petrova, E.G., Strekalovsky, A.S.: The quadratic-linear bilevel problems solving via nonconvex constraint problems. Int. J. Biomed. Softw Comput. Hum. Sci. (Special Issue on Variational Inequality and Combinatorial Problems) **18**(1), 63–67 (2011)
27. Strekalovsky, A.S.: On problem of global extremum. Proc. USSR Acad. Sci. **292**(5), 1062–1066 (1987)
28. Strekalovsky, A.S.: Elements of Nonconvex Optimization. Nauka Publication, Novosibirsk (2003) (in Russian)
29. Strekalovsky, A.S.: On the minimization of the difference of convex functions on a feasible set. Comput. Math. Math. Phys. **43**(3), 399–409 (2003)
30. Strekalovsky, A.S.: Minimizing sequences in problems with D.C. constraints. Comput. Math. Math. Phys. **45**(3), 418–429 (2005)
31. Strekalovsky, A.S., Gruzdeva, T.V.: Local search in problems with nonconvex constraints. Comput. Math. Math. Phys. **47**(3), 397–413 (2007)
32. Strekalovsky, A.S., Orlov, A.V.: A new approach to nonconvex optimization. Numer. Methods Program. **8**(2), 11–27 (2007) (in Russian)

33. Strekalovsky, A.S., Orlov, A.V.: Bimatrix Games and Bilinear Programming. Physmathlit, Moscow (2007) (in Russian)
34. Strekalovsky A.S., Petrova, E.G., Mazurkevich, E.O.: On numerical solving a linear problem of complementarity. Comput. Math. Math. Phys. **49**(8), 1318–1331 (2009)
35. Strekalovsky, A.S., Orlov, A.V., Malyshev, A.V.: Local search in a quadratic-linear bilevel programming problem. Numer. Anal. Appl. **3**(1), 59–70 (2010)
36. Strekalovsky, A.S., Orlov, A.V., Malyshev, A.V.: Numerical solution of a class of bilevel programming problems. Numer. Anal. Appl. **3**(2), 165–173 (2010)
37. Strekalovsky, A.S., Orlov, A.V., Malyshev, A.V.: On computational search for optimistic solutions in bilevel problems. J. Glob. Optim. **48**, 159–172 (2010)
38. Trefethen, L.N., Bau, D.: Numerical Linear Algebra. SIAM, Philadelphia (1997)
39. Tuy, H.: D.C. optimization: theory, methods and algorithms. In: Horst, R., Pardalos, P.M. (eds.) Handbook of Global Optimization, pp. 149–216. Kluwer, Dordrecht (1995)
40. Vasilyev, F.P.: Optimization Methods. Factorial Press, Moscow (2002) (in Russian)