A Simple But Effective Canonical Dual Theory Unified Algorithm for Global Optimization

Jiapu Zhang

Graduate School of Information Technology and Mathematical Sciences,
University of Ballarat, Victoria, Australia

Emails: j.zhang@ballarat.edu.au, jiapu zhang@hotmail.com, Phone: 61-423 487 360

Abstract

Numerical global optimization methods are often very time consuming and could not be applied for high-dimensional nonconvex/nonsmooth optimization problems. Due to the nonconvexity/nonsmoothness, directly solving the primal problems sometimes is very difficult. This paper presents a very simple but very effective canonical duality theory (CDT) unified global optimization algorithm. This algorithm has convergence is proved in this paper. More important, for this CDT-unified algorithm, numerous numerical computational results show that it is very powerful not only for solving low-dimensional but also for solving high-dimensional nonconvex/nonsmooth optimization problems, and the global optimal solutions can be easily and elegantly got with zero dual gap.

1 Introduction

In recent years large-scale global optimization (GO) problems have drawn considerable attention. These problems have many applications, in particular in data mining, computational biology/chemistry. Numerical methods for GO are often very time consum-ing and could not be applied for high-dimensional nonconvex/nonsmooth optimization problems. Due to the nonconvexity/nonsmoothness, directly solving the primal prob-lems sometimes is very difficult; however, in this paper, their prime-dual Gao-Strang complementary problems enable us to elegantly and easily not only get the global optimal solutions of primal problems but also of canonical dual problems. The canonical duality theory (CDT) was originally developed for nonconvex/nonsmooth mechanics [7]. It is now realized that this potentially powerful theory can be used for solving a large class of nonconvex/nonsmooth GO problems [3, 9, 10, 12, 13, 14, 15] with applications to data mining clustering, sensor network problems, and molecular distance geometry problem [20], etc.. In this paper, a CDT-unified GO algorithm is presented. According to the CDT, the proof of the convergent theorem for the algorithm is very easy and clear. High-dimensional nonconvex/nonsmooth GO examples (such as the minimization problem of Rosenbrock function) are tested by the new algorithm and numerical results pleasantly show that the new algorithm has a great promise for solving some GO problems.

2 The Algorithm and its Convergence

The algorithm is established from CDT [7]. In this paper we will solve the following nonconvex GO problem [18, 16]:

$$(\mathcal{P}) \qquad \min_{x}\{P(x) = V(\Lambda(x)) - F(x) | x \in \mathcal{X}_a\}, \tag{1}$$

where $\mathcal{X}_a \subset \mathbb{R}^n$ is a feasible space, $V(\Lambda(x))$ is not necessarily convex with respect to x and it is a so-called canonical function satisfying

$$V^*(\varsigma) = sta\{\langle \xi; \varsigma \rangle - V(\xi) | \xi \in \mathcal{V}\} : \mathcal{V}^* \to \mathbb{R}, \tag{2}$$

$$\varsigma = \nabla V(\xi) \Leftrightarrow \xi \in \nabla V^*(\varsigma) \Leftrightarrow V(\xi) + V^*(\varsigma) = \langle \xi; \varsigma \rangle, \tag{3}$$

Λ is a geometrically admissible (objective) mapping from the feasible space $\mathcal{X}_a$ into a canonical measure space $\mathcal{V}$, $F(x) = \langle x, f \rangle - \frac{1}{2}\langle x, Ax \rangle$, $\langle *, * \rangle$ denotes a bilinear form in $\mathbb{R}^n \times \mathbb{R}^n, f \in \mathbb{R}^n, A = A^T \in \mathbb{R}^{n \times n}$ are given, and $\langle *; * \rangle$ represents a bilinear form which puts $\mathcal{V}$ and $\mathcal{V}^*$ in duality. By (2)-(3), the nonconvex function $P(x)$ in (1) can be rewritten as

$$\Xi(x, \varsigma) = \langle \Lambda(x); \varsigma \rangle - V^*(\varsigma) - F(x). \tag{4}$$

In real applications, the objective operator Λ is usually a quadratic measure over a given a field [7, 18] and in finite space this quadratic measure can be written as [8, 18, 16]:

$$\Lambda(x) = \left\{ \frac{1}{2} x^T B_k x + b_k^T x + c_k = \xi_k, k = 1, 2, \ldots, m \right\} = \xi : \mathbb{R}^n \to \mathbb{R}^m, \tag{5}$$

where $B_k \in \mathbb{R}^{n \times n}$ is a symmetrical matrix, $b_k \in \mathbb{R}^n$, $c_k \in \mathbb{R}^1$ for each $k = 1, 2, \ldots, m$. Denote

$$G(\varsigma) = A + \sum_{k=1}^{m} \varsigma_k B_k : \mathbb{R}^m \to \mathbb{R}^{n \times n}, \tag{6}$$

$$d(\varsigma) = f - \sum_{k=1}^{m} \varsigma_k b_k : \mathbb{R}^m \to \mathbb{R}^n, \tag{7}$$

and define

$$S_a = \{\varsigma \in \mathbb{R}^m | d(\varsigma) \in Col(G(\varsigma))\}, \tag{8}$$
$$S_a^+ = \{\varsigma \in S_a | G(\varsigma) \succeq 0\}, \tag{9}$$

where $Col(G(\varsigma))$ is the space spanned by the column of $G(\varsigma)$, and $G(\varsigma) \succeq 0$ stands for $G(\varsigma)$ is a semi-positive definite matrix (if $detG(\varsigma) = 0, G(\varsigma)^{-1}$ should be understood as the generalized inverse [8, 18, 16]).

The following algorithm is designed:

Algorithm 1 - *A canonical dual theory algorithm.*

Step 1. Call the subroutine Algorithm 2, or simply the *Matlab's fsolve* program if dimension of problems are less than a few thousands, to solve differential equations $\Xi(x, \varsigma)' = 0$.

Step 2. Output the roots $\bar{\varsigma}, \bar{x}$.

Step 3. Check whether $\bar{\varsigma} \in S_a$ are in S_a^+; if so, then

Step 4. Pick up the $\bar{x}$s that are satisfying $G(\bar{\varsigma})\bar{x} = d(\bar{\varsigma})$.

Theorem 1 *(Convergence of the Algorithm) $\bar{x}$ generated by Algorithm 1 is the global optimal solution of ($\mathcal{P}$).*

Proof. By Theorem 3 of [18] and its reference [6] we know that $\bar{x}$ generated by Algorithm 1 is a global optimal solution of ($\mathcal{P}$). ■

The powerful of this simple algorithm can be easily demonstrated by easily and effectively solving the following benchmark test problems of GO, even calculated by hands on paper.

Example 1. (Two-dimensional Rosenbrock function) $P(x) = (x_1 - 1)^2 + 100(x_2 - x_1^2)^2$.
Solution. $\Xi(x, \varsigma) = (x_1 - 1)^2 + (x_1^2 - x_2)\varsigma - \frac{1}{400}\varsigma^2$ and $S_a^+ = \{\varsigma \in \mathbb{R}^1 | \varsigma > -1\}$ are got. Let $\Xi(x, \varsigma)' = 0$ a critical point $(\bar{x}_1, \bar{x}_2, \bar{\varsigma}) = (1, 1, 0)$ is got. We easily know $\bar{\varsigma} = 0 \in S_a^+$ and satisfying $G(\bar{\varsigma})\bar{x} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = f = d(\bar{\varsigma})$. Thus, $(\bar{x}_1, \bar{x}_2) = (1, 1)$ is a global minimum of Rosenbrock function.

Example 2. (Two-dimensional De Jong function) $P(x) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2 + (x_2 - 1)^2$.
Solution. $\Xi(x, \varsigma) = (x_1 - 1)^2 + (x_2 - 1)^2 + (x_1^2 - x_2)\varsigma - \frac{1}{400}\varsigma^2$ and $S_a^+ = \{\varsigma \in \mathbb{R}^1 | \varsigma > -1\}$ are got. Let $\Xi(x, \varsigma)' = 0$ a critical point $(\bar{x}_1, \bar{x}_2, \bar{\varsigma}) = (1, 1, 0)$ is got. We easily know $\bar{\varsigma} = 0 \in S_a^+$ and satisfying $G(\bar{\varsigma})\bar{x} = f = d(\bar{\varsigma})$. Thus, $(\bar{x}_1, \bar{x}_2) = (1, 1)$ is a global minimal

solution.

Example 3. (Colville function) $P(x) = 100(x_2 - x_1^2)^2 + 90(x_3^2 - x_4)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 10.1((x_2 - 1)^2 + (x_4 - 1)^2) + 19.8(x_2 - 1)(x_4 - 1)$.
Solution. Let $\Xi(x, \varsigma)' = 0$ and a unique critical point $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{\varsigma}_1, \bar{\varsigma}_2) = (1, 1, 1, 1, 0, 0)$ is got such that $(\bar{\varsigma}_1, \bar{\varsigma}_2) \in S_a^+$ and $G(\bar{\varsigma})\bar{x} = f = d(\bar{\varsigma})$. Thus, $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = (1, 1, 1, 1)$ is a global minimal solution.

Example 4. (Two-dimensional Zakharov function) $P(x) = x_1^2 + x_2^2 + (0.5x_1 + x_2)^2 + (0.5x_1 + x_2)^4$.
Solution. Let $\Xi(x, \varsigma)' = 0$ and a unique critical point $(\bar{x}_1, \bar{x}_2, \bar{\varsigma}) = (0, 0, 0)$ is got such that $\bar{\varsigma} \in S_a^+$, satisfying the formula $G(\bar{\varsigma})\bar{x} = \begin{pmatrix} 5/2 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = f = d(\bar{\varsigma})$. Thus, (0,0) is the global optimal solution.

Example 5. (Four-dimensional Powell function) $P(x) = (x_1 - 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - x_3)^4 + 10(x_1 - x_4)^4$.
Solution. Let $\Xi(x, \varsigma)' = 0$ and a unique critical point $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4, \bar{\varsigma}_1, \bar{\varsigma}_2) = (0, 0, 0, 0, 0, 0)$ is got such that $(\bar{\varsigma}_1, \bar{\varsigma}_2) \in S_a^+$ and $G(\bar{\varsigma})\bar{x} = f = d(\bar{\varsigma})$. Thus, $(\bar{x}_1, \bar{x}_2, \bar{x}_3, \bar{x}_4) = (0, 0, 0, 0)$ is a global minimal solution.

Example 6. (Double Well function) $P(x) = \frac{1}{2}(\frac{1}{2}x^2 - 2)^2 - \frac{1}{2}x$.
Solution. $\Xi(x, \varsigma) = (\frac{1}{2}x^2 - 2)\varsigma - \frac{1}{2}\varsigma^2 - \frac{1}{2}x$ and $S_a^+ = \{\varsigma \in \mathbb{R}^1 | \varsigma > 0\}$ are got. Let $\Xi(x, \varsigma)' = 0$ and three critical points of $\Xi(x, \varsigma)$: $(\bar{x}^1, \bar{\varsigma}^1) = (2.11491, 0.236417)$, $(\bar{x}^2, \bar{\varsigma}^2) = (-1.86081, -0.268701)$, $(\bar{x}^3, \bar{\varsigma}^3) = (-0.254102, -1.96772)$ are got, and a unique point $\bar{\varsigma}^1$ in S_a^+ is got. Thus, $\bar{x}^1 = 2.11491$ is the global minimal solution of the Double Well function. This can also be illuminated in Figure 1.

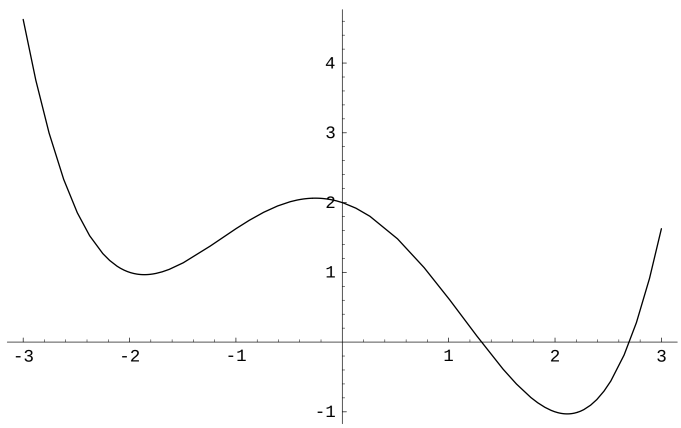

Figure 1: The prime double-well function.

The canonical dual problem of Double Well function is:

$$\max_{\varsigma > 0} -\frac{1}{8\varsigma} - \frac{1}{2}\varsigma^2 - 2\varsigma, \tag{10}$$

which is equal to

$$\min_{\varsigma>0} \frac{1}{8\varsigma} + \frac{1}{2}\varsigma^2 + 2\varsigma \tag{11}$$

that can be transferred into (but not equal to (because of the constant of ς_0)) (this idea is communicated from Wu C.Z.)

$$\min_{t,\varsigma} t \quad subject \quad to \begin{pmatrix} \varsigma & 0.5/\sqrt{2} \\ 0.5/\sqrt{2} & t - \frac{1}{2}\varsigma_0^T\varsigma - 2\varsigma \end{pmatrix} \succeq 0 \tag{12}$$

and solved by Semi-Definite Programming (SDP) package SeDuMi 1_3 [24] (this idea is communicated from Wu C.Z. In this whole paper, this is the single idea of other researchers). However, after 6 circles in all 33 iterations SeDuMi 1_3 died at the solution 0.2585, which is still very far from the real global optimal solution 0.236417 of the dual problem (10). This means the popular SDP package SeDuMi 1_3 has a poor performance for Double Well function minimization problem, compared with our Algorithm 1. But other strategies to solve the following dual problem of (1) are sought:

$$\min_{G(\varsigma)\succeq 0} P^d(\varsigma) \quad = \quad \frac{1}{2}d(\varsigma)^T G(\varsigma)^{-1} d(\varsigma) + \sum_{k=1}^{m}\left(\frac{1}{2}\alpha_k^{-1}\varsigma_k^2 - c_k\varsigma_k\right); \tag{13}$$

for example, one strategy is to replace (13) by the following quadratic semidefinite programming problem (QSDP)

$$\min_{t,\varsigma} \quad t + \frac{1}{2}\alpha^{-1}\varsigma^T\varsigma + +\varsigma^T c \quad subject \quad to \begin{pmatrix} G(\varsigma) & \frac{1}{\sqrt{2}}d(\varsigma) \\ \frac{1}{\sqrt{2}}d(\varsigma)^T & t \end{pmatrix} \succeq 0, \tag{14}$$

where α, c, ς are vectors, and some known QSDP packages can solve (14) for low-dimensional problems and for high-dimensional problems we can design QSDP algorithms by ourselves. This is one direction of algorithm design for CDT. Another direction is to efficiently and effectively get the roots of $\Xi(x,\varsigma)' = 0$ for Algorithm 1, i.e. to solve the $m+n$ quadratic equations (15) given as follows.

For high-dimensional nonconvex GO problems, the finite element method (FEM)-based [25] subroutine of Algorithm 1 is well designed as follows. *Step 1* of Algorithm 1 is to find the roots of $\Xi(x,\varsigma)' = 0$, i.e. the following $m+n$ quadratic equations:

$$\begin{cases} \frac{1}{2}x^T B_k x + b_k^T x + c_k = \alpha_k^{-1}\varsigma_k, k = 1,2,\ldots,m & (by \quad \Xi(x,\varsigma)'_\varsigma = 0), \\ G(\varsigma)x = d(\varsigma), \quad i.e. \quad (A + \sum_{k=1}^m \varsigma_k B_k)x = f - \sum_{k=1}^m \varsigma_k b_k & (by \quad \Xi(x,\varsigma)'_x = 0), \end{cases} \tag{15}$$

where $\alpha_k, k = 1,2,\ldots,m$ are the coefficients in $V(\Lambda(x)) = \sum_{k=1}^m \frac{1}{2}\alpha_k(\frac{1}{2}x^T B_k x + b_k^T x + c_k)^2$.

Algorithm 2 - *A subroutine finding roots of (15). E.g. the finite element discretized* $\Xi(x,\varsigma)'$ *method if the dimension of the problem is large than a few thousands.*

Firstly the minimization problem for Rosenbrock function for (15) is tested. In global optimization, the nonconvex minimization problem of Rosenbrock function is a benchmark test problem that is extensively used to test the performance of optimization algorithms and approaches. The global minimum is inside a long, deep, narrow, parabolic/banana shaped flat valley. The shallow global minimum is inside a deeply curved valley. To find the valley and to converge to the global minimum is difficult. Detailed introduction to the difficulty of the problem and the excellence as the testing problem of GO can be referred to [19].

Minimizing the Rosenbrock function:

$$
\begin{aligned}
\min P(x) = &\quad \sum_{k=1}^{n-1}[100(x_k^2 - x_{k+1})^2 + (x_k - 1)^2] : x \in \mathbb{R}^n && (16)\\
= &\quad n - 1 + \sum_{k=1}^{n-1} \frac{1}{2}\alpha_k (x_k^2 - x_{k+1})^2 + \sum_{k=1}^{n-1} x_k^2 - \sum_{k=1}^{n-1} 2x_k && (17)\\
= &\quad n - 1 + \sum_{k=1}^{n-1} \frac{1}{2}\alpha_k \left(\frac{1}{2}x^T B_k x + b_k^T x + c_k\right)^2 + \frac{1}{2}x^T A x - x^T f, && (18)
\end{aligned}
$$

where $\alpha_k = 200$, $B_k = \begin{pmatrix} 0 & & & & & & & \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & 0 & & & & \\ & & & & 2 & & & \\ & & & & & 0 & & \\ & & & & & & \cdot & \\ & & & & & & & \cdot \\ & & & & & & & & 0 \end{pmatrix}_{n\times n}$, $b_k = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ -1 \\ 0 \\ \cdot \\ 0 \end{pmatrix}_{n\times 1}$, $c_k = 0, k = 1, \ldots, n-1$, $A = \begin{pmatrix} 2 & & & & & & & \\ & \cdot & & & & & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & \cdot & & & \\ & & & & & \cdot & & \\ & & & & & & \cdot & \\ & & & & & & & 2 & \\ & & & & & & & & 0 \end{pmatrix}_{n\times n}$ $f = \begin{pmatrix} 2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ 2 \\ 0 \end{pmatrix}_{n\times 1}$, $m = n - 1$, and $S_a^+ = \{\varsigma_k > -1, k = 1, 2, \ldots, n-1\}$. For Rosenbrock function, (15) is written as:

$$
\begin{cases}
x_k^2 - x_{k+1} = 0.005\varsigma_k, k = 1, 2, \ldots, n-1, \\
2(1 + \varsigma_{k+1})x_{k+1} = 2 + \varsigma_k, k = 1, 2, \ldots, n-3, \\
\varsigma_{n-1} = 0, \\
(1 + \varsigma_1)x_1 = 1, \\
2x_{n-1} = 2 + \varsigma_{n-2},
\end{cases} \tag{19}
$$

(19) is solved by Matlab's *fsolve* on the Intel(R) Celeron(R) CPU 900@2.20GHz Windows Vista$^{\mathrm{TM}}$ Home Basic personal notebook computer. The initial solution for $(x;\varsigma)$ is set as $(3,\ldots,3;2,\ldots,2)$ and the numerical computational results are shown in Table 1. We may see in Table 1 that the global minimal solution $(1,\ldots,1)$ for (1) can be easily and accurately got by Algorithm 1, at the same time the global maximal solution $(0,\ldots,0)$ for the (13) over S_a^+ can be easily and accurately got by Algorithm 1 directly too. The smart Matlab's *fsolve* does not output any other solution in S/S_a^+.

Table 1: Results of numerical computations

Dimension n	Iterations	Calls of Equations (19)	prime and dual opt sln		CPU time
3	6	42	$(1,\ldots,1)$	$(0,\ldots,0)$	0.0468
4	6	56	$(1,\ldots,1)$	$(0,\ldots,0)$	0.0468
5	6	70	$(1,\ldots,1)$	$(0,\ldots,0)$	0.0624
6	7	96	$(1,\ldots,1)$	$(0,\ldots,0)$	0.0780
7	7	112	$(1,\ldots,1)$	$(0,\ldots,0)$	0.1248
8	7	128	$(1,\ldots,1)$	$(0,\ldots,0)$	0.1248
9	7	144	$(1,\ldots,1)$	$(0,\ldots,0)$	0.0936
10	7	160	$(1,\ldots,1)$	$(0,\ldots,0)$	0.1404
20	7	320	$(1,\ldots,1)$	$(0,\ldots,0)$	0.2808
30	8	540	$(1,\ldots,1)$	$(0,\ldots,0)$	0.4368
40	8	720	$(1,\ldots,1)$	$(0,\ldots,0)$	0.6084
50	8	900	$(1,\ldots,1)$	$(0,\ldots,0)$	0.8736
60	8	1,080	$(1,\ldots,1)$	$(0,\ldots,0)$	1.0920
70	8	1,260	$(1,\ldots,1)$	$(0,\ldots,0)$	1.2948
80	8	1,440	$(1,\ldots,1)$	$(0,\ldots,0)$	1.6536
90	8	1,620	$(1,\ldots,1)$	$(0,\ldots,0)$	1.9344
100	8	1,800	$(1,\ldots,1)$	$(0,\ldots,0)$	2.2464
200	9	4,000	$(1,\ldots,1)$	$(0,\ldots,0)$	7.7844
300	9	6,000	$(1,\ldots,1)$	$(0,\ldots,0)$	15.8497
400	9	8,000	$(1,\ldots,1)$	$(0,\ldots,0)$	27.4250
500	9	10,000	$(1,\ldots,1)$	$(0,\ldots,0)$	42.4947
600	9	12,000	$(1,\ldots,1)$	$(0,\ldots,0)$	61.2616
700	9	14,000	$(1,\ldots,1)$	$(0,\ldots,0)$	88.2342
800	9	16,000	$(1,\ldots,1)$	$(0,\ldots,0)$	118.3736
900	10	19,800	$(1,\ldots,1)$	$(0,\ldots,0)$	170.8367
1,000	10	22,000	$(1,\ldots,1)$	$(0,\ldots,0)$	221.8334
2,000	10	44,000	$(1,\ldots,1)$	$(0,\ldots,0)$	1491.9

Chemical database clustering problem. Algorithm 1 is applied to solve the real chemical database clustering problem (34)-(35) of [27], i.e.

$$\min P(Y_1,Y_2,\ldots,Y_n) = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\frac{1}{2}\alpha_{ij}\left(||Y_i-Y_j||^2-d_{ij}^2\right)^2 \tag{20}$$

$$= \sum_{i,j=1,i<j}^{n}\frac{1}{2}\alpha_{ij}\left((y_{i1}-y_{j1})^2+(y_{i2}-y_{j2})^2-d_{ij}^2\right)^2, \tag{21}$$

where $\alpha_{ij} = \frac{1}{2d_{ij}^4}$ if $d_{ij}^4 \geq 10^{-12}$, $\alpha_{ij} = \frac{1}{2}$ if $d_{ij}^4 < 10^{-12}$, for all $i, j = 1, 2, \ldots, n$ (d_{ij} is calculated from the original n by 9 dataset [27]) and $Y_i \in \mathbb{R}^2, i = 1, 2, \ldots, n$. The *Gao-Strang generalized complementary function* for (21) is:

$$\Xi(Y, \varsigma) = \sum_{i,j=1,i<j}^{n} \left((||Y_i - Y_j||^2 - d_{ij}^2)\varsigma_{ij} - \frac{1}{2\alpha_{ij}}\varsigma_{ij}^2 \right) \tag{22}$$

$$= \sum_{i,j=1,i<j}^{n} \left(\left((y_{i1} - y_{j1})^2 + (y_{i2} - y_{j2})^2 - d_{ij}^2\right) \varsigma_{ij} - \frac{1}{2\alpha_{ij}}\varsigma_{ij}^2 \right). \tag{23}$$

Choose the SVD-reduced initial solution $Y_i^0, i = 1, 2, \ldots, n$ [27, 29] and calculate $\varsigma_{ij}^0 = \alpha_{ij}(||Y_i^0 - Y_j^0||^2 - d_{ij}^2)$ as the initial solution for ς_{ij}^0 in solving the following quadratic nonlinear equations of $\Xi(Y, \varsigma)'$=0:

$$\begin{cases} 2(y_{i1} - y_{j1})\varsigma_{ij} = 0, i = 1, 2, \ldots, n-1, j = i+1, \ldots, n, \\ -2(y_{i1} - y_{j1})\varsigma_{ij} = 0, i = 1, 2, \ldots, n-1, j = i+1, \ldots, n, \\ 2(y_{i2} - y_{j2})\varsigma_{ij} = 0, i = 1, 2, \ldots, n-1, j = i+1, \ldots, n, \\ -2(y_{i2} - y_{j2})\varsigma_{ij} = 0, i = 1, 2, \ldots, n-1, j = i+1, \ldots, n, \\ (y_{i1} - y_{j1})^2 + (y_{i2} - y_{j2})^2 - d_{ij}^2 - \frac{1}{\alpha_{ij}}\varsigma_{ij} = 0, i = 1, 2, \ldots, n-1, j = i+1, \ldots, n. \end{cases} \tag{24}$$

The global optimal prime or dual solution is got and the comparison of Algorithm 1 with the SD (steepest descent) method, CG (conjugate gradient) method, BFGS (approximated Newton) method, NR (Newton-Raphson) method, and TN (truncated Newton) method T-IHN (truncated incomplete Hessian Newton) method, etc can be made. The equations (24) are illuminated to make readers to easily understand the equations. When $n = 2$, i.e. $i = 1, j = 2$ there are the following 5 quadratic equations with 5 variables (with initial value $(y_{11}^0, y_{12}^0, y_{21}^0, y_{22}^0; \varsigma_{12}^0) = (331.5590, -188.4908, 364.6889, -158.1010; 0)$):

$$\begin{cases} 2(y_{11} - y_{21})\varsigma_{12} = 0, \\ -2(y_{11} - y_{21})\varsigma_{12} = 0, \\ 2(y_{12} - y_{22})\varsigma_{12} = 0, \\ -2(y_{12} - y_{22})\varsigma_{12} = 0, \\ (y_{11} - y_{21})^2 + (y_{12} - y_{22})^2 - 15701874.4205486\varsigma_{12} - 2801.95239257813 = 0, \end{cases} \tag{25}$$

In (24), different database of [29] has different n. Numerical computational results of solving (24) will be updated.

For CDT, there are two research directions for its algorithm design. One is to design the CDT algorithm to solve (13); for example, one strategy is to design the quadratic semidefinite programming (QSDP) algorithm to solve (14). Another research direction is to design the CDT algorithm to solve the special m + n quadratic (nonlinear) equations, with m+n variables, (15); for example, Newton-type iteration algorithms, gradient-type iteration algorithms, trust-region-type iteration algorithms, nonlinear finite-element-type algorithms, Hamiltonian system symplectic-type algorithms, etc are good strategy to solve (15). Researchers may design a powerful CDT algorithm along these two directions to solve (13) and (15) respectively.

3 Advantages of Solving the Dual Problem Than Solving the Prime Problem

In this section, we still use the benchmark GO test problem, minimizing the Rosenbrock function, to illuminate some advantages of solving the canonical dual problem (13) compared with directly solving the prime problem (1). By the CDT [7, 16], the canonical dual problem of (16) is:

$$\max_{\varsigma>-1} P^d(\varsigma) = n - 1 - \sum_{k=1}^{n-1} \left[\frac{(\varsigma_{k-1}+2)^2}{4(\varsigma_k+1)} + \frac{1}{400}\varsigma_k^2 \right], \tag{26}$$

where $\varsigma_0 = 0$. (26) is solved by the Discrete Gradient (DG) method [1], a local search optimization solver for nonconvex and/or nonsmooth optimization problems, and the numerical computational results are listed in Tables 2-3 (where seed1 is the initial solution $(x^0;\varsigma^0) = (3,3,\ldots,3;-2/3,-2/3,\ldots,-2/3,0)$ and seed2 is the initial solution $(x^0;\varsigma^0) = (100,100,\ldots,100;100,100,\ldots,100,0)$).

In Tables 2-3, we may see that the dual problem (26) can be elegantly, easily, quickly and accurately solved to get its objective function value 0.00000000, compared with the prime problem (16). We may also see that (26) is convenient for MPI (Message Passing Interface) parallel computation. The successfully tested MPI code is followed:

broadcast $n-1$

call MPI_BCAST ($n-1$,1,MPI_INTEGER, 0, MPI_COMM_WORLD , ierr)

check for quit signal

if ($n-1$ .le. 0) goto 30

calculate every partials

sum = 0.0d0

do 20 i = myid+1, $n-1$, numprocs

if ($i-1$ .eq. 0) then $\varsigma(0)$=0

sum = sum $+(\varsigma(i-1)+2.0)/(4*(\varsigma(i)+1.0))+(1.0/400.0)*\varsigma(i)**2$

20 continue

f = sum

collect all the partial sums

call MPI_REDUCE (f,objf,1,MPI_DOUBLE_PRECISION, MPI_SUM, 0,
& MPI_COMM_WORLD, ierr)

30 node 0 (i.e. myid = 0) prints the sums = objf

Table 2: Results of numerical experiments for seed1

Dimension n	Iterations		Function calls		Objective function value	
	Prime	Dual	Prime	Dual	Prime	Dual
2	120	24	2843	28	0.00001073	0.00000000
3	422	26	8996	137	0.00401438	0.00000000
4	3737	35	48352	202	0.00615273	0.00000000
5*	335	34	10179	399	3.96077434	0.00000000
6*	2375	44	43770	868	4.00635895	0.00000000
7*	1223	53	28009	1625	4.09419146	0.00000000
8	2160	55	46792	2100	0.01246714	0.00000000
9	2692	51	61017	2526	0.01397307	0.00000000
10	4444	63	91470	3979	0.01055630	0.00000000
20	3042	55	140924	10084	0.00940077	0.00000000
30	2321	58	133980	20515	0.01075478	0.00000000
40	1659	60	173795	26818	0.01227866	0.00000000
50	2032	57	219233	36459	0.01264147	0.00000000
60	1966	61	260701	50495	0.01048188	0.00000000
70	1876	56	272919	52545	0.01531147	0.00000000
80	1405	61	195156	59684	0.01594730	0.00000000
90	2142	61	371963	71320	0.01055831	0.00000000
100	2676	60	510722	70208	0.01125514	0.00000000
200	1395	61	653604	188589	0.01115318	0.00000000
300	1368	60	882760	235163	0.01574873	0.00000000
400	2085	66	1869675	301805	0.00928066	0.00000000
500	1155	59	1394240	358938	0.01168440	0.00000000
600	1226	63	1808285	451817	0.00918730	0.00000000
700	1557	60	2134359	559378	0.01257100	0.00000000
800	1398	61	2098062	522726	0.01442714	0.00000000
900	716	65	1904187	763449	0.01074534	0.00000000
1000	1825	61	3598608	681509	0.00897202	0.00000000
2000	257	62	2087277	1455472	0.00937219	0.00000000
3000	3221	60	20642543	2714296	0.01250373	0.00000000
4000*	679	60	7581502	3659292	4.11193171	0.00000000

This shows the advantages of solving the dual problem (13) than solving the prime problem (1). We may use maximizing the canonical dual function of Colville function (see Example 3) to illuminate this point again.

It is known that directly solving the minimizing of Colville function is difficult. But, it is very easy to solving the following canonical dual problem:

$$
\begin{aligned}
&\max_{\varsigma>-1} P^d(\varsigma) = 42 - \frac{1}{400}\varsigma_1^2 - \frac{1}{360}\varsigma_2^2 \\
&-\frac{1}{2}\begin{pmatrix} 2 \\ 40+\varsigma_1 \\ 2 \\ 40+\varsigma_2 \end{pmatrix}^T \begin{pmatrix} 2+2\varsigma_1 & 0 & 0 & 0 \\ 0 & 20.2 & 0 & 19.8 \\ 0 & 0 & 2+2\varsigma_2 & 0 \\ 0 & 19.8 & 0 & 20.2 \end{pmatrix}^+ \begin{pmatrix} 2 \\ 40+\varsigma_1 \\ 2 \\ 40+\varsigma_2 \end{pmatrix},
\end{aligned}
$$

Table 3: Results of numerical experiments for seed2

Dimension n	Iterations		Function calls		Objective function value	
	Prime	Dual	Prime	Dual	Prime	Dual
2	10013	24	227521	28	47.23824896	0.00000000
3	144	32	4869	235	96.49814330	0.00000000
4	144	81	5279	938	82.46230602	0.00000000
5	148	137	5682	1768	94.19254867	0.00000000
6	154	166*	6238	2590	88.84382963	0.00000000
7	159	179*	7097	3288	237.63078399	0.00000000
8	165	202*	7502	4300	238.41126013	0.00000000
9	153	206*	7137	5083	84.54205412	0.00000000
10	162	216*	7491	5920	83.23094398	0.00000000
20	225	285*	19111	17458	83.94779152	0.00000000
30	216	301*	20939	28543*	156.95838274	0.00000000
40	163	291*	19775	40444*	83.30960344	0.00000000
50	158	298*	33269	51888*	85.93091895	0.00000000
60	158	312*	34094	61767*	89.07412094	0.00000000
70	162	284*	35436	69865*	92.45725362	0.00000000
80	209	297*	35607	89127*	157.69955825	0.00000000
90	227	294*	60398	98748*	82.44035053	0.00000000
100	202	290*	57792	102796*	81.94595276	0.00000000
200	1826	262	436413	189293	83.77165551	0.00000000
300	195	259*	169238	261320*	152.95671738	0.00000000
400	195	278*	212104	375816*	82.49253919	0.00000000
500	190	297*	331637	522695*	82.40170647	0.00000000
600	292	303*	431092	559068*	150.15456693	0.00000000
700	189	275*	383735	758631*	89.14575473	0.00000000
800	198	270*	429674	701053*	84.50538257	0.00000000
900	198	280*	416150	867398*	85.32757049	0.00000000
1000	193	283*	445326	930761*	89.48369379	0.00000000
2000	232	310*	1123240	2030104*	84.26810981	0.00000000

which is a simple problem of maximizing a concave function over a convex set: by watching Figure 2, $\bar{\varsigma} = (0, 0)$ is easily known as the global optimal solution.

Thus, the following optimization algorithm designed to solve the canonical dual problem (14) is very necessary.

Algorithm 3 - *An optimization algorithm to solve the Quadratic Semidefinite Programming (14).*

In Example 6, (11) is equal to

$$
\begin{aligned}
\min_{t,\varsigma\in\mathbb{R}} \quad & \frac{1}{2}\begin{pmatrix} t \\ \varsigma \end{pmatrix}^T \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ \varsigma \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix}^T \begin{pmatrix} t \\ \varsigma \end{pmatrix} \\
s.t. \quad & \begin{pmatrix} \varsigma & 0.5/\sqrt{2} \\ 0.5/\sqrt{2} & t \end{pmatrix} \succeq 0,
\end{aligned}
$$

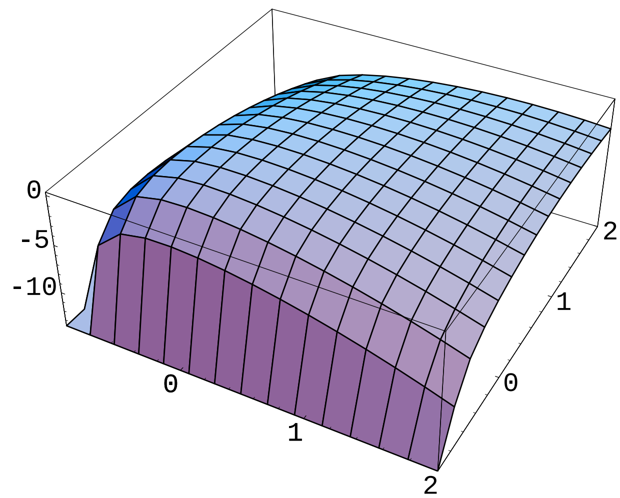

Figure 2: Canonical dual Colville function on $S_a^+ = \{\varsigma > -1\}$

which can solved by Algorithm 3.

References

[1] Bagirov, A.M., Karasozen, B., Sezer, M.: Discrete gradient method: a derivative free method for nonsmooth optimization. J. Opt. Theor. Appl. 137: 317-334 (2008)

[2] Bagirov, A.M., Rubinov, A.M., Zhang, J.P.: Local optimization method with global multidimensional search. J. Glob. Optim. 32: 161–179 (2005)

[3] Fang, S.C., Gao, D.Y., Sheu, R.L., Wu, S.Y.: Canonical dual approach to solving 0-1 quadratic programming problems. J. Ind. Manag. Opt. 4 (1): 125–142 (2008)

[4] Feng, K.: Numerical Computational Algorithms. National Defence Industrial Press (1978) (in Chinese)

[5] Feng, K., Qin M.Z.: Hamiltonian Symplectic System. Zhejiang Science and Technology Press (2003) (in Chinese)

[6] Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlienar beam theory. Appl. Mech. Rev. 50 (11), S64-S71 (1997)

[7] Gao, D.Y.: Duality Principles in Nonconvex Systems: Theory, Methods and Applications. Kluwer Academic Publishers, Dordrecht/Boston/London (2000)

[8] Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. J. Glob. Optim. 17 (1), 127-160 (2000)

[9] Gao, D.Y.: Perfect duality theory and complete set of solutions to a class of global optimization problems. Opt. 52 (4–5): 467–493 (2003)

[10] Gao, D.Y.: Canonical duality theory and solutions to constrained nonconvex quadratic programming. J. Glob. Opt. 29: 377–399, (2004)

[11] Gao, D.Y.: Complementary principle, algorithm, and complete solution to phase transitions in solids governed by Landau-Ginzburg equation. Math. Mech. Solids 9: 61-79 (2004)

[12] Gao, D.Y.: Sufficient conditions and perfect duality in nonconvexs minimization with inequality. J. Ind. Manag. Opt. 1 (1): 53-63 (2005)

[13] Gao, D.Y.: Complete solutions and extremality criteria to polynomial optimization problems, J. Glob. Opt. 35 : 131-143 (2006)

[14] Gao, D.Y.: Solutions and optimality criteria to box constrained nonconvex minimization problems, J. Ind. Manag. Opt. 3(2), 293-304 (2007)

[15] Gao, D.Y.: Advances in canonical duality theory with applications to global optimization. Proceedings Foundations of Computer-Aided Process Operations (FOCAPO 2008), June 29July 2, 2008, Cambridge, MA (2008)

[16] Gao, D.Y., Ruan, N., Pardalos, P.: Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization (2010)

[17] Gao, D.Y., Ruan, N., Zhang, J.: Canonical dual solutions to a global optimization problem in database analysis (2011)

[18] Gao, D.Y., Wu, C.Z.: On the triality theory in global optimization, J. Glob. Opt. (2011) accepted, arXiv:1104.2970v1

[19] Gao, D.Y., Zhang, J.P.: Canonical duality theory for solving minimization problem of Rosenbrock function, J. Glob. Opt. (2011) in revision

[20] Gao, D.Y., Zhang, J.P.: A novel canonical dual computational approach for prion AGAAAAGA amyloid fibril molecular modeling. J. Theor. Biol.: in revision (2011)

[21] Ghosh, R., Rubinov, A.M., Zhang, J.P.: Optimization approach for clustering datasets with weights. Optim. Method Softw. 20 (2) 335–351 (2005)

[22] Li, Q.Y., Mo, Z.Z., Qi, L.Q.: Numerical Solutions for the Systems of Nonlinear Equations Beijing: Seience Press (1987) (in Chinese)

[23] Lu, C., Wang, Z.B., Xing, W.X., Fang, S.C.: Extended canonical duality and conic programming for solving 0-1 quadratic programming problem, J. Ind. Manag. Optim. 6 (4): 779-793 (2010)

[24] Lu, W.S.: Use SeDuMi to Solve LP, SDP and SCOP Problems: Remarks and Examples (December 2, 2009 Version):
http://www.ece.uvic.ca/∼wslu/Talk/SeDuMi-Remarks.pdf

[25] Santosa, H.A.F.A, Gao, D.Y.: Canonical dual finite element method for solving post-buckling problems of a large deformation elastic beam. Int. J. Nonlinear Mech. (to appear) (2011)

[26] Sun, K., Tu, S.K., Gao, D.Y., Xu, L.: Canonical dual approach to binary factor analysis. The Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation 2009, Paraty, Brazil, March 15–18, 2009.

[27] Xie, D.X. and Ni, Q.: An incomplete Hessian Newton minimization method and its application in a chemical database problem. Comput. Optim. Appl. 44: 467–485 (2009)

[28] Xie, D.X., Singh, S.B., Fluder, E.M. and Schlick, T.: Principal component analysis combined with truncated-Newton minimization for dimensionality reduction of chemical databases. Math. Program. 95: 161–185 (2003)

[29] Xie, D.X., Tropsha, A. and Schlick, T.: An efficient projection protocol for chemical databases: singular value decomposition combined with truncated-newton minimization. J. Chem. Inf. Comput. Sci. 40: 167–177 (2000)