Canonical Duality Theory for Solving Minimization Problem of Rosenbrock Function

David Yang Gao · Jiapu Zhang

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

Abstract This paper presents a *canonical duality theory* for solving nonconvex minimization problem of Rosenbrock function. Extensive numerical results show that this benchmark test problem can be solved precisely and efficiently to obtain global optimal solutions.

Keywords global optimization · canonical duality · NP-hard problems · triality

1 Introduction

Nonconvex minimization problem of Rosenbrock function, introduced in [15], is a benchmark test problem in global optimization that has been used extensively to test performance of optimization algorithms and numerical approaches. The global minimizer of this function is located in a long, deep, narrow, parabolic/banana shaped flat valley (Figure 1).

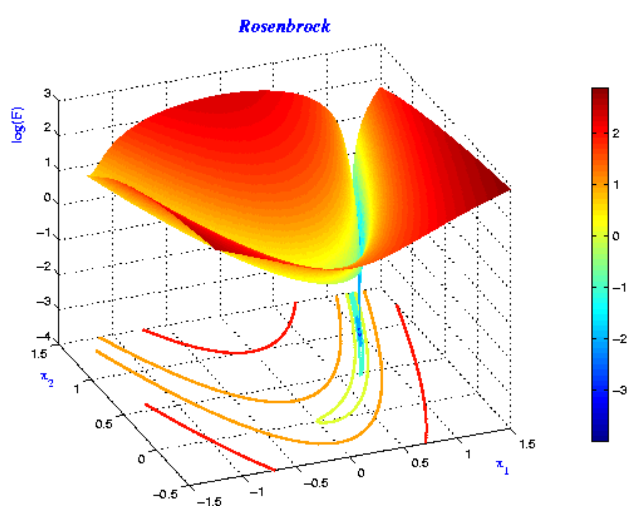

Figure 1: 2-dimensional Rosenbrock function (www2.imm.dtu.dk/courses/02610/)

Although to find this valley is trivial for most cases, to accurately locate the global optimal solution is very difficult by almost all gradient-type methods and some derivative-free methods. Due to the nonconvexity, it can be easily tested that if the initial point is chosen to be $(3, 3, \ldots, 3)$, direct algorithms are always trapped into a local minimizer for problems with dimensions $n = 5 \sim 7$ as well as $n \geq 4000$; if the initial point is chosen at $(100, 100, \ldots, 100)$, iterations will be stopped at a local min with the objective function value > 47.23824896 even for a two-dimensional problem. This paper will show that by the canonical duality theory, this well-known benchmark problem can be solved efficiently in an elegant way.

The *canonical duality theory* was originally developed in nonconvex/nonsmooth mechanics [9]. It is now realized that this potentially powerful theory can be used for solving a large class of nonconvex/nonsmooth/discrete problems [10, 12]. In this short research note, we will first show the nonconvex minimization problem of Rosenbrock function can be reformulated as a canonical dual problem (with zero duality gap) and the critical point of the Rosenbrock function can be analytically expressed in terms of its canonical dual solutions. Both global and local extremal solutions can be identified by the triality theorem. Extensive numerical examples and discussion are presented in the last section.

2 Primal Problem and Its Canonical Dual

The primal problem is

$$(\mathcal{P}): \quad \min \left\{ P(\mathbf{x}) = \sum_{i=1}^{n-1} \left[(x_i - 1)^2 + \frac{1}{2}\alpha (x_{i+1} - x_i^2)^2 \right] \;\; | \;\; \mathbf{x} \in \mathcal{X} \right\}, \tag{1}$$

where $\mathbf{x} = \{x_i\} \in \mathcal{X} = \mathbb{R}^n$ is a real unknown vector, $\alpha = 2N$ and N is a given real number. Clearly, this is a nonconvex minimization problem which could have multiple local minimizers.

In order to use the canonical duality theory for solving this nonconvex problem, we need to define a *geometrically admissible* canonical measure

$$\boldsymbol{\xi} = \{\xi_i\} = \{x_i^2 - x_{i+1}\} \in \mathcal{E}_a \subset \mathbb{R}^{n-1}. \tag{2}$$

The canonical function $V : \mathcal{E}_a \to \mathbb{R}$ can be defined by

$$V(\boldsymbol{\xi}) = \sum_{j=1}^{n-1} \frac{1}{2}\alpha\xi_j^2, \tag{3}$$

which is a convex function. The canonical dual variable $\boldsymbol{\varsigma} = \boldsymbol{\xi}^*$ can be defined uniquely by

$$\boldsymbol{\varsigma} = \{\varsigma_j\} = \nabla V(\boldsymbol{\xi}) = \{\alpha\xi_j\}. \tag{4}$$

Therefore, by the Legendre transformation, the conjugate function $V^* : \mathcal{S} = \mathbb{R}^{n-1} \to \mathbb{R}$ is obtained as

$$V^*(\boldsymbol{\varsigma}) = \mathrm{sta}\{\boldsymbol{\xi}^T\boldsymbol{\varsigma} - V(\boldsymbol{\xi}) \mid \boldsymbol{\xi} \in \mathcal{E}_a\} = \sum_{j=1}^{n-1} \frac{1}{2}\alpha^{-1}\varsigma_j^2. \tag{5}$$

Replacing $\sum_{i=1}^{n-1} \frac{1}{2}\alpha(x_{i+1} - x_i^2)^2$ by the Legendre equality $V(\Lambda(\mathbf{x})) = \Lambda(\mathbf{x})^T\boldsymbol{\varsigma} - V^*(\boldsymbol{\varsigma})$, the total complementary function $\Xi : \mathcal{X} \times \mathcal{S} \to \mathbb{R}$ is given by

$$\Xi(\mathbf{x}, \boldsymbol{\varsigma}) = \sum_{i=1}^{n-1} \left[(x_i - 1)^2 + \varsigma_i(x_i^2 - x_{i+1}) - \frac{1}{2}\alpha^{-1}\varsigma_i^2 \right]. \tag{6}$$

Let $\delta^\sharp$ and $\delta^\flat$ be shifting operators such that $\delta^\sharp\varsigma_i = \varsigma_{i+1}$ and $\delta^\flat\varsigma_i = \varsigma_{i-1}$. We define $\delta^\flat\varsigma_1 = 0$. Then on the canonical dual feasible space

$$\mathcal{S}_a = \{\boldsymbol{\varsigma} \in \mathcal{S} | \ \varsigma_i + 1 \neq 0 \ \ \forall i = 1, \ldots, n-2, \ \varsigma_{n-1} = 0\}, \tag{7}$$

the canonical dual can be obtained by

$$P^d(\boldsymbol{\varsigma}) = \mathrm{sta}\{\Xi(\mathbf{x}, \boldsymbol{\varsigma}) | \ \mathbf{x} \in \mathcal{X}\} = n - 1 - \sum_{i=1}^{n-1} \left[\frac{(\delta^\flat\varsigma_i + 2)^2}{4(\varsigma_i + 1)} + \frac{1}{2}\alpha^{-1}\varsigma_i^2 \right]. \tag{8}$$

Based on the *complementary-dual principle* proposed in the canonical duality theory (see [**?**]), we have the following result.

Theorem 1 *If $\bar{\boldsymbol{\varsigma}}$ is a critical point of $P^d(\boldsymbol{\varsigma})$, then the vector $\bar{\mathbf{x}} = \{\bar{x}_i\}$ defined by*

$$\bar{x}_i = \frac{\delta^\flat\bar{\varsigma}_i + 2}{2(\bar{\varsigma}_i + 1)}, \ \ i = 1, \cdots, n-1, \ \ \bar{x}_n = \bar{x}_{n-1}^2 \tag{9}$$

is a critical point of $P(\mathbf{x})$ and

$$P(\bar{\mathbf{x}}) = \Xi(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}}) = P^d(\bar{\boldsymbol{\varsigma}}). \tag{10}$$

This theorem presents actually an "analytic" solution form to the Rosenbrock function, i.e. the critical point of the Rosenbrock function must be in the form of (9) for each dual solution $\bar{\varsigma}$. The first version of this analytical solution form was presented in nonconvex variational problems in phase transitions and finite deformation mechanics [5, 6, 7]. The extremality of the analytical solution is governed by the so-called *triality theory*. Let

$$\mathcal{S}_a^+ = \{\boldsymbol{\varsigma} \in \mathcal{S}_a | \ \varsigma_i + 1 > 0 \ \ \forall i = 1, \dots, n-1\}, \tag{11}$$

we have the following theorem:

Theorem 2 *Suppose that $\bar{\boldsymbol{\varsigma}}$ is a critical point of $P^d(\boldsymbol{\varsigma})$ and the vector $\bar{\mathbf{x}} = \{\bar{x}_i\}$ is defined by Theorem 1.*

If $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^+$, then $\bar{\boldsymbol{\varsigma}}$ is a global maximal solution to the canonical dual problem on $\mathcal{S}_a^+$, i.e.,

$$(\mathcal{P}^d_+): \quad \max\{P^d(\boldsymbol{\varsigma}) \mid \boldsymbol{\varsigma} \in \mathcal{S}_a^+\}, \tag{12}$$

the vector $\bar{\mathbf{x}}$ is a global minimal to the primal problem, and

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}} P(\mathbf{x}) = \max_{\boldsymbol{\varsigma} \in \mathcal{S}_a^+} P^d(\boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\varsigma}}). \tag{13}$$

Theorem 2 shows that the canonical dual problem $(\mathcal{P}^d_+)$ provides a global optimal solution to the nonconvex primal problem. Since $(\mathcal{P}^d_+)$ is a concave maximization problem over a convex space which can be solved easily. This theorem is actually a special application of Gao and Strang's general result on global minimizer in in nonconvex analysis [13].

By introducing

$$\mathcal{S}_a^- = \mathcal{S}/\mathcal{S}_a^+ = \{\boldsymbol{\varsigma} \in \mathbb{R}^{n-1} | \ \varsigma_i + 1 < 0 \ \ \forall i = 1, \dots, n-1\}, \tag{14}$$

recently the triality theory (see [14]) leads to the following theorem.

Theorem 3 *Suppose that $\bar{\boldsymbol{\varsigma}}$ is a critical point of $P^d(\boldsymbol{\varsigma})$ and the vector $\bar{\mathbf{x}} = \{\bar{x}_i\}$ is defined by Theorem 1.*

If $\bar{\boldsymbol{\varsigma}} \in \mathcal{S}_a^-$, then on a neighborhood $\mathcal{X}_o \times \mathcal{S}_o \subset \mathcal{X} \times \mathcal{S}_a^-$ of $(\bar{\mathbf{x}}, \bar{\boldsymbol{\varsigma}})$, we have either

$$P(\bar{\mathbf{x}}) = \min_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \min_{\boldsymbol{\varsigma} \in \mathcal{S}_o} P^d(\boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\varsigma}}), \tag{15}$$

or

$$P(\bar{\mathbf{x}}) = \max_{\mathbf{x} \in \mathcal{X}_o} P(\mathbf{x}) = \max_{\boldsymbol{\varsigma} \in \mathcal{S}_o} P^d(\boldsymbol{\varsigma}) = P^d(\bar{\boldsymbol{\varsigma}}), \tag{16}$$

The proof of this Theorem can be derived from the recent paper by Gao and Wu [14]. By the fact that the canonical dual function is a d.c. function (difference of convex functions) on $\mathcal{S}_a^-$, the double-min duality (15) can be used for finding the biggest local minimizer of the Rosenbrock function $P(\mathbf{x})$, while the double-max duality (16) can be used for finding the biggest local maximizer of $P(\mathbf{x})$. In physics and material sciences, this pair of biggest local extremal points play important roles in phase transitions.

Because $\varsigma_{n-1} = 0$, we may know that $\mathcal{S}_a^-$ is an empty set. Thus, by Theorem 3 in this paper we cannot find a local maximizer or minimizer on $\mathcal{S}_a^-$ or its subset for $P^d(\bar{\boldsymbol{\varsigma}})$.

3 Numerical Examples and Discussion

$(\mathcal{P})$ and $(\mathcal{P}_+^d)$ will be solved by the discrete gradient (DG) method ([2]), which is a local search optimization solver for nonconvex and/or nonsmooth optimization problems. In two dimensional space, Rosenbrock function has a long ravine with very steep walls and flat bottom; "because of the curved flat valley the optimization is zig-zagging slowly with small stepsizes towards the minimum" (en.wikipedia.org/wiki/Gradient_descent). This means any gradient method may fail to minimize the Rosenbrock function even from 2 dimensions. The DG method is a derivative-free method which can be applied for miminizing/maximizing Rosenbrock function and its dual. Numerical experiments have been carried out in Intel(R) Celeron(R) CPU 900@2.20GHz Windows Vista$^{\mathrm{TM}}$ Home Basic personal notebook computer.

We try N=100 (when N=10 we find the numerical results are similar to $N = 100$), with the dimensions n=2∼10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000. We first set $(3, 3, \ldots, 3)$ (called seed1) as the initial solution for $(\mathcal{P})$ (usually the feasible solution space is a box constrained between -2.048 and 2.048 [1, 16, 17]). Numerical results (Table 1) show that to solve the primal problem $(\mathcal{P})$, the DG method can easily and quickly get approximate global minimum solution to $\bar{\mathbf{x}} = (1, 1, \ldots, 1)$ with the approximate global optimal values at $P(\bar{\mathbf{x}}) = 0$, except for n=5∼7, 4000, where the DG method can only get a local minimum solution $\bar{\mathbf{x}} = (-1, 1, \ldots, 1)$ with $P(\bar{\mathbf{x}}) = 4$. Then we let $\mathbf{x}_0 = (100, 100, \ldots, 100)$ (called seed2) be the initial solution for $(\mathcal{P})$, searched in the intervals $-500 \le x_i \le 500, i = 1, 2, \ldots, n$. We find that the DG method was trapped into local optimal solutions but not getting any global minimum at all, even from a 2 dimensional problem (see Table 2), its objective function value is 47.23824896. However, from Table 2 we can see that by the same DG method, the global maximum of the dual problem can be obtained very elegantly.

For $(\mathcal{P}_+^d)$, the corresponding dimensions are 1∼9, 19, 29, 39, 49, 59, 69, 79, 89, 99, 199, 299, 399, 499, 599, 699, 799, 899, 999, 1999, 2999, 3999. The initial solution is set as $\boldsymbol{\varsigma}_0 = (-2/3, -2/3, \ldots, -2/3, 0)$ (called seed1), the constraints $\boldsymbol{\varsigma}_i + 1 > 0, i = 1, 2, \ldots, n-1$ were penalized into the objective function; by $\Xi(\mathbf{x}, \boldsymbol{\varsigma})'_{x_n} = 0$ of formula (6), we can set the values of the last variable ς_{n-1} always being 0 (> -1). With these numerical computation settings, the DG method can easily and quickly solve all these test problems to accurately get a global maximizer $\bar{\boldsymbol{\varsigma}} = (0, 0, ..., 0)$ with the optimal value $P^d(\bar{\boldsymbol{\varsigma}}) = 0$ (Table 1). By the fact that the canonical dual problem $(\mathcal{P}_+^d)$ is a concave maximization over a convex open space, the DG method was not trapped into any local optimal solution. But, for the nonconvex primal problem $(\mathcal{P})$ in dimensions n=5∼7 and 4000, the DG method was trapped into local minimizer $\bar{\mathbf{x}} = (-1, 1, \ldots, 1)$. If we set the initial solution as $\boldsymbol{\varsigma}_0 = (100, 100, \ldots, 100, 0)$ (called seed2) and repeat the calculations, our numerical results (Table 2) show again that the canonical dual problem can be easily and quickly solved by the DG method to accurately get the global maximizer $\bar{\boldsymbol{\varsigma}} = (0, 0, ..., 0)$ with the optimal solution $P^d(\bar{\boldsymbol{\varsigma}}) = 0$ for dimensions $n = 1 \sim 9, 19, 29, 39, 49, 59, 69, 79, 89, 99, 199, 299, 399, 499, 599, 699, 799, 899, 999, 1999$.

The comparisons between $(\mathcal{P})$ and $(\mathcal{P}_+^d)$ in view of total number of iterations and total number of objective function evaluations (i.e. function calls) are listed in Tables 1-2. Compared with $(\mathcal{P}_+^d)$, the approximate global and local optimal solutions and their optimal objective function values of $(\mathcal{P})$ are not accurate, and even cannot be obtained if the initial iteration is set to be $\mathbf{x}_0 = (100, 100, \ldots, 100)$. In Table 1, we can see that the total number of iterations and function calls for $(\mathcal{P})$ are always greater than those for $(\mathcal{P}_+^d)$. This means that $(\mathcal{P}_+^d)$ costs less computer calculations than $(\mathcal{P})$, though $(\mathcal{P}_+^d)$ still can get accurate global optimal solutions and the global optimal objective function value. The initial solutions $\mathbf{x}_0 = (100, 100, \ldots, 100)$ and $\boldsymbol{\varsigma}_0 = (100, 100, \ldots, 100, 0)$ respectively for $(\mathcal{P})$ and $(\mathcal{P}_+^d)$ are not practical for real numerical tests so that the total number of iterations and function calls of $(\mathcal{P})$ are sometimes less than those of $(\mathcal{P}_+^d)$. Regarding the CPU times for solving $(\mathcal{P}_+^d)$ with $n = 4000$, the largest CPU time for seed1 is 206.3581 seconds (i.e. 3.4393 minutes).

Example 1. Let $n = 4$ (four dimensions). The global minimizer is known to be

Table 1: Results of numerical experiments for $(\mathcal{P})$ and $(\mathcal{P}_+^d)$: $N = 100$, seed1

Dimension n	Iterations		Function calls		Objective function value	
	$(\mathcal{P})$	$(\mathcal{P}_+^d)$	$(\mathcal{P})$	$(\mathcal{P}_+^d)$	$(\mathcal{P})$	$(\mathcal{P}_+^d)$
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

Table 2: Results of numerical experiments for $(\mathcal{P})$ and $(\mathcal{P}_+^d)$: $N = 100$, seed2

Dimension n	Iterations		Function calls		Objective function value	
	$(\mathcal{P})$	$(\mathcal{P}_+^d)$	$(\mathcal{P})$	$(\mathcal{P}_+^d)$	$(\mathcal{P})$	$(\mathcal{P}_+^d)$
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

$\bar{\mathbf{x}} = (1, 1, 1, 1)$ and $P(\bar{\mathbf{x}}) = 0$.

Solution: By using the DG method for both primal problem $(\mathcal{P})$ and its canonical dual $(\mathcal{P}_+^d)$, we have the numerical solutions

$$\bar{\mathbf{x}} = (1.0166873133, 1.0337174892, 1.0687306765, 1.1425101552), \quad P(\bar{\mathbf{x}}) = 0.00615273,$$

$$\bar{\boldsymbol{\varsigma}} = (0.0000000119, 0.0000000000, 0.0000000000), \quad P_+^d(\bar{\boldsymbol{\varsigma}}) = 0.00000000.$$

This shows that the canonical dual problem provides more accurate solution.

Example 2. For dimension $n = 5$, the Rosenbrock function has exactly two minima, one is the global optimal solution $(1, 1, 1, 1, 1)$ with global optimal minimum value 0, and another minimum is a local minimum near $(-1, 1, 1, 1, 1)$ with local optimal minimum value 4.

Solution: By the DG method, the primal solution is

$$\bar{\mathbf{x}} = (-0.9856129203, 0.9814803343, 0.9682775584, 0.9398661046, 0.8840549028)$$

with $P(\bar{\mathbf{x}}) = 3.96077434$. Clearly, this is a local minimizer. While the canonical dual problem produces accurately a global optimal solution

$$\bar{\boldsymbol{\varsigma}} = (0.0000004388, 0.0000006036, 0.0000000000, 0.0000000000), \quad P_+^d(\bar{\boldsymbol{\varsigma}}) = 0.$$

Example 3. For $n = 6$ (six dimensions), the Rosenbrock function has exactly two minima, i.e., the global optimal solution

$$\bar{\mathbf{x}}_1 = (1, 1, 1, 1, 1, 1), \quad P(\bar{\mathbf{x}}_1) = 0,$$

and local minimal solution

$$\bar{\mathbf{x}}_2 = (-1, 1, 1, 1, 1, 1), \quad P(\bar{\mathbf{x}}_2) = 4.$$

Solution: To solve the primal problem directly, the DG method can only provide local solution

$$\bar{\mathbf{x}} = (-0.9970726441, 1.0041582933, 1.0133158817, 1.0292928527, 1.0607123926, 1.1258344785)$$

with $P(\bar{\mathbf{x}}) = 4.00635895$. For the canonical dual problem, the DG method produces

$$\bar{\boldsymbol{\varsigma}} = (0.0000001747, -0.0000000559, 0.0000005919, 0.0000000000, 0.0000000000),$$

$$P_+^d(\bar{\boldsymbol{\varsigma}}) = 0.$$

Example 4. Similarly, if $n = 7$, the test problem has the same global optimal solution

$$\bar{\mathbf{x}}_1 = (1,1,1,1,1,1,1), \quad P(\bar{\mathbf{x}}_1) = 0$$

and the local minimal solution

$$\bar{\mathbf{x}}_2 = (-1,1,1,1,1,1,1), \quad P(\bar{\mathbf{x}}_2) = 4.$$

Solution: By the DG method, we have

$$\begin{aligned} \bar{\mathbf{x}} &= (-1.0003403494, 1.0106728675, 1.0264433859, 1.0561180077, \\ &\quad 1.1168007274, 1.2483026410, 1.5594822181), \\ P(\bar{\mathbf{x}}) &= 4.09419146, \end{aligned}$$

$$\begin{aligned} \bar{\boldsymbol{\varsigma}} &= (-0.0000001431, -0.0000011147, -0.0000010643, -0.0000003284, \\ &\quad 0.0000000000, 0.0000000000), \\ P_+^d(\bar{\boldsymbol{\varsigma}}) &= 0. \end{aligned}$$

This shows again that the DG iterations for solving the primal problem is trapped to a local min.

Example 5. Now we let $n = 4000$. The Rosenbrock function has many minima. The global optimal solution is still $\bar{\mathbf{x}}_1 = (1, \ldots, 1)$ with $P(\bar{\mathbf{x}}) = 0$. One of local minima is nearby the point $\bar{\mathbf{x}}_2 = (-1, 1, \ldots, 1)$ with $P(\bar{\mathbf{x}}_2) = 4$.

Solution: Again, by the DG method, the primal iteration is trapped at

$$\bar{\mathbf{x}} = (-0.9932861006, 0.9966510741, \ldots, 1.3122885708, 1.7233744896), \quad P(\bar{\mathbf{x}}) = 4.11193171.$$

The conical dual solution is

$$\begin{aligned} \bar{\boldsymbol{\varsigma}} &= (-0.0000000314, -0.0000000040, -0.0000000437, \ldots, \\ &\quad -0.0000000281, 0.0000000008, -0.0000000214, 0.0000000000, 0.0000000000), \end{aligned}$$

which produce precisely the optimal value $P_+^d(\bar{\boldsymbol{\varsigma}}) = 0$. Indeed, as long as $n \geq 5$, the DG method is always trapped into the local minimizer $\bar{\mathbf{x}} = (-1, 1, \ldots, 1)$ if the initial solution is set to be $\mathbf{x}_0 = (-1.0005, 1.0005, \ldots, 1.0005)$.

It is worth to note that both $P(\mathbf{x})$ and $P^d(\boldsymbol{\varsigma})$ are the sum of $n-1$ items. This is convenient for MPI (Message Passing Interface) parallel computations. We may broadcast (MPI_Bcast) the sum to $n-1$ processes, each process calculates one item, and at last all the partials are reduced (MPI_Reduce) onto one process to get the sum. Thus on Tambo machines of VLSCI (http://www.vlsci.unimelb.edu.au) we should be able to successfully solve (1) and (12) with at least 3.2767×10^7 variables if setting the maximal variables for the DG method to be 4000 (though the DG method and its parallelization version ([3]) can solve optimization problems with more than 4000 variables). 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

sum = sum $+(x(i)-1.0)**2+100.0*(x(i)**2-x(i+1))**2$

20 continue (for $P(\mathbf{x})$)

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 (for $P^d(\boldsymbol{\varsigma})$)

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

4 Conclusion

This research note demonstrates a powerful application of the *canonical duality theory* for solving the nonconvex minimization problem of Rosenbrock function. Extensive numerical computations show that by using the same DG method, the canonical dual problem can be easily solved to produce global solutions.

Acknowledgments:

This research is supported by US Air Force Office of Scientific Research under the grant AFOSR FA9550-10-1-0487. The authors thank Professor Kok Lay Teo (Curtin University, Australia) for his helpful comments and Associate Professor Adil M. Bagirov (Ballarat University, Australia) for his FORTRAN codes of his discrete gradient method.

References

[1] Ai, Y.S., Liu, P.C., Zheng, T.Y.: Adaptive hybrid global inversion algorithm. Science in China (Series D). **41 (2)**, 137–143 (1998)

[2] Bagirov, A.M.: Continuous subdifferential approximations and their applications. J. Math. Sci. **15**, 2567–2609 (2003)

[3] Beliakov, G., Monsalve Tobon, J.E., Bagirov, A.M.: Parallelization of the discrete gradient method of non-smooth optimization and its applications. Computational Science – ICCS 2003 Lecture Notes in Computer Science. **2659/2003**, 701–711 (2003)

[4] Bouvry, P., Arbab, F., Seredynski, F.: Distributed evolutionary optimization, in Manifold: Rosenbrock's function case study. Inf. Sci. **122**, 141–159 (2000)

[5] Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Applied Mechanics Reviews Vol. **50 (11)**, S64–S71 (1997)

[6] Gao, D.Y.: General analytic solutions and complementary variational principles for large deformation nonsmooth mechanics. Meccanica **34**, 169–198 (1999)

[7] Gao, D.Y.: Analytic solution and triality theory for nonconvex and nonsmooth variational problems with applications. Nonlinear Analysis **42 (7)**, 1161–1193 (2000)

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

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

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

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

[12] 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)

[13] Gao, D.Y. and Strang, G.: Geometric nonlinearity: Potential energy, complementary energy, and the gap function. Quart. Appl. Math. XLVII(3), 487–504 (1989)

[14] Gao, D.Y. and Wu, C.Z.: On the triality theory in global optimization. J. Global Optimization (2010)

[15] Rosenbrock, H.H.: An automatic method for finding the greatest or least value of a function. The Computer Journal **3**: 175–184 (1960)

[16] Xin, B., Chen, J., Pan, F.: Problem difficulty analysis for particle swarm optimization: deception and modality. 2009 Proceedings of the 1st ACM/SIGEVO Summit on Genetic and Evolutionary Computation 623–630 (2009)

[17] Yu, H.F., Wang, D.W.: Research on food-chain algorithm and its parameters. Front. Electr. Electron. Eng. China **3(4)**, 394-398 (2008)