A SEQUENTIAL CONVEX PROGRAM METHOD TO DC PROGRAM WITH JOINT CHANCE CONSTRAINTS

XIANTAO XIAO  
School of Mathematical Sciences  
Dalian University of Technology  
Dalian 116023, China

JIAN GU  
School of Sciences  
Dalian Ocean University  
Dalian 116023, China

LIWEI ZHANG  
School of Mathematical Sciences  
Dalian University of Technology  
Dalian 116023, China

SHAOWU ZHANG  
School of Computer Science and Technology  
Dalian University of Technology  
Dalian 116023, China

(Communicated by Xiaooqi Yang)

Abstract. In this paper, we consider a DC (difference of convex) programming problem with joint chance constraints (JCCDCP). We propose a DC function to approximate the constrained function and a corresponding DC program \((P_\varepsilon)\) to approximate the JCCDCP. Under some mild assumptions, we show that the solution of Problem \((P_\varepsilon)\) converges to the solution of JCCDCP when \(\varepsilon \downarrow 0\). A sequential convex program method is constructed to solve the Problem \((P_\varepsilon)\). At each iteration a convex program is solved based on the Monte Carlo method, and the generated optimal sequence is proved to converge to the stationary point of Problem \((P_\varepsilon)\).

1. Introduction. In this paper, we consider the following optimization problem:

\[
\begin{align*}
\text{minimize} & \quad f_1(x) - f_2(x) \\
\text{subject to} & \quad c_1(x, \xi) \leq 0, \ldots, c_m(x, \xi) \leq 0, \\
& \quad x \in \mathcal{X},
\end{align*}
\]

where \(\mathcal{X} \subseteq \mathbb{R}^n\) is a compact convex set, \(\xi \in \mathbb{R}^k\) is a parameter vector, \(f_i : \mathbb{R}^n \to \mathbb{R}, i = 1,2\) and \(c_j : \mathbb{R}^{n+k} \to \mathbb{R}, j = 1, \ldots, m\) are real-valued convex functions. Let \(f(x) := f_1(x) - f_2(x)\), the function \(f(x)\) is called DC (difference of convex) function, and Problem (1) is a DC programming problem. The DC program (1) has received a lot of interests, since many complicated optimization problems can be...
transformed into DC program. For instance, optimization problem with complementarity constraints, quadratic programming with quadratic constraints [1], low rank optimization problem [4], etc. See [7] for a good survey on applications, mathematical properties and numerical methods of DC programming.

In this paper, we assume that the parameter $\xi$ of Problem (1) is a random vector (which is very common in practical applications), then the problem (1) is a stochastic DC programming problem, which seems impossible to be solved, even a feasible point is very difficult to obtain. In order to overcome the difficulty, we relax the constraints and consider the following problem

\[
\begin{align*}
\text{(JCCDCP)} \\
\text{minimize} & \quad f_1(x) - f_2(x) \\
\text{subject to} & \quad \Pr\{c_1(x, \xi) \leq 0, \ldots, c_m(x, \xi) \leq 0\} \geq \alpha, \\
& \quad x \in \mathcal{X}.
\end{align*}
\]

The constraint $\Pr\{c_1(x, \xi) \leq 0, \ldots, c_m(x, \xi) \leq 0\} \geq \alpha$ is called a joint chance constraint, which requires all $m$ constrains be satisfied simultaneously with a probability greater than $\alpha$, where $0 < \alpha < 1$ is usually set as 0.9, 0.95 or 0.99. The joint chance constrained problems were first introduced in [2], since then they have been extensively studied by stochastic optimization scholars.

For the following discussion, we define

\[ c(x, \xi) := \max\{c_1(x, \xi), \ldots, c_m(x, \xi)\}, \]

and

\[ p(x) := 1 - \Pr\{c(x, \xi) \leq 0\} = \Pr\{c(x, \xi) > 0\}, \]

then the joint chance constraint is in the form of

\[ p(x) \leq 1 - \alpha. \tag{2} \]

Note that $p(x) = \Pr\{c(x, \xi) > 0\} = E[1_{(0,+\infty)}(c(x, \xi))]$, where $1_{(0,+\infty)}(\cdot)$ is the indicator function defined by

\[ 1_{(0,+\infty)}(z) = \begin{cases} 
1, & \text{if } z > 0, \\
0, & \text{otherwise}.
\end{cases} \]

Usually, the function $p(x)$ is not convex and has no closed form. To solve the joint chance constrained problem, many conservative approximations of $p(x)$ have been proposed. Among all convex approximations, the conditional value-at-risk (CVaR) approximation [9] is proved to be the best, which uses

\[ \text{CVaR}_\alpha(c(x, \xi)) \leq 0 \]

to be a conservative approximation of constraint (2), where

\[ \text{CVaR}_\alpha(z) := \inf_{t \in \mathbb{R}} \{(1 - \alpha)^{-1}E[z - t]^+ + t\} \quad \text{and} \quad [z]^+ := \max(z, 0). \]

The constraint $\text{CVaR}_\alpha(c(x, \xi)) \leq 0$ can be rewritten as

\[ \inf_{t > 0} \{t^{-1}E[t + c(x, \xi)]^+_+\} \leq 1 - \alpha. \]

It is shown that the CVaR approximation is actually using $\frac{1}{t}[t + z]^+$ to approximate $1_{(0,+\infty)}(z)$. From Figure (1) we can see that, although CVaR approximation is the best convex approximation of $p(x)$, however, it is not a good choice since the approximation becomes terrible as $z \to +\infty$. Hong, Yang and Zhang [6] proposed a better but nonconvex approximation, they used a DC function

\[ \frac{1}{t}[t + z]^+ - \frac{1}{t}[z]^+. \]
to approximate $1_{(0, +\infty)}(z)$. From Figure (1), when $t > 0$ is very small, the approximation is much better than CVaR approximation. But the drawback is also obvious, it is not convex.

By using the DC approximation function to approximate $p(x)$, the JCCDCP can be approximated by a constrained problem with both object function and constraint function are DC functions. In this paper, we show that, under some mild assumptions, the optimal solution of the DC approximation problem is an approximate solution of JCCDCP, when the parameter $t$ is close to 0. To solve the DC programming problem, we propose a sequential convex program method.

The study in this paper gives a deep insight into the work of Hong, Yang and Zhang [6] and extends the results in [6] to a more general form. Compared to the work in the literature, the problem we consider is a stochastic DC programming problem which is more general and more complicated. The assumptions used in [6] are very strict, especially for Assumption 3 “$c(x, \xi)$ is differentiable with respect to $x$ w.p.1”, which in fact reduces the joint chance constraint to a single chance constraint. In this paper, we give a more relaxed version than Assumption 3, and remove some other assumptions.

The rest of this paper is organized as follows. In Section 2, we review the DC approximation of the joint chance constraints, and develop a good approximation of JCCDCP, denoted by ($P_\varepsilon$). In Section 3, we propose a sequential convex program method to solve Problem ($P_\varepsilon$). The numerical results are reported in Section 4, followed by the conclusions in Section 5.

2. Approximation of JCCDCP. In this section, we propose an approximation of Problem (P) in the form of

$$\begin{align*}
\text{(P)} \quad &\text{minimize} \quad f_1(x) - f_2(x) \\
&\text{subject to} \quad p(x) \leq 1 - \alpha \\
&\quad x \in \mathcal{X},
\end{align*}$$

which is equivalent to JCCDCP.

2.1. DC approximation. Define

$$\phi(z, t) := [z + t]_+, \quad \pi(z, t) := \frac{1}{t} [\phi(z, t) - \phi(z, 0)].$$

As mentioned previously, $\pi(z, t)$ is a better approximation of $1_{(0, +\infty)}(z)$ than $\frac{1}{t} \phi(z, t)$. Let

$$\begin{align*}
g_1(x, t) &:= \mathbb{E}[t + c(x, \xi)]_+ = \mathbb{E}[\phi(c(x, \xi), t)], \\
g_2(x) &:= g_1(x, 0) = \mathbb{E}[\phi(c(x, \xi), 0)].
\end{align*}$$

Note that $g_1(x, t)$ and $g_2(x)$ are both convex in $x$. Define

$$\tilde{p}(x, t) := \mathbb{E}[\pi(c(x, \xi), t)] = \frac{1}{t} [g_1(x, t) - g_2(x)],$$

Figure 1. CVaR and DC approximation
then for any $t > 0$, $\hat{p}(x,t)$ is a conservative approximation of $p(x) = \mathbb{E}[1_{(0,\infty)}(c(x, \xi))]$.

Before the discussion, we give the following assumptions.

**Assumption 2.1.** The support of $\xi$, denoted by $\Xi$, is a closed subset of $\mathbb{R}^k$. For any $\xi \in \Xi$, $f_1(x), f_2(x)$ and $c_i(x, \xi), i = 1, \ldots, m$, are continuously differentiable for any $x \in \mathcal{O}$ where $\mathcal{O}$ is a bounded open set such that $\mathcal{X} \subset \mathcal{O}$.

**Assumption 2.2.** There exists a random variable $K_i$ with $\mathbb{E}[K_i] < +\infty$ such that

$$|c_i(x_1, \xi) - c_i(x_2, \xi)| \leq K_i \|x_1 - x_2\|$$

for any $x_1, x_2 \in \mathcal{O}$ and any $i = 1, 2, \ldots, m$.

Assumption 2.2 implies that $c_i(x, \xi), i = 1, \ldots, m$ are Lipschitz continuous.

Let $F(t, x) = \Pr\{c(x, \xi) \leq t\}$ be the cumulative distribution function of $c(x, \xi)$.

**Assumption 2.3.** There exists $\delta > 0$ such that $F(t, x)$ is continuous on $(-\delta, \delta) \times \mathcal{O}$.

**Assumption 2.4.** Let $\Omega_0 = \{x \in \mathcal{X} : p(x) \leq 1 - \alpha\}$ and $\Omega_0^I = \{x \in \mathcal{X} : p(x) < 1 - \alpha\}$. Then $\Omega_0 = c(\Omega_0^I)$.

Assumption 2.4 is a kind of constraint qualification. It is implied by the Slater condition when $\Omega_0$ is convex.

**Assumption 2.5.** $g_2(\cdot)$ is differentiable on $\mathcal{O}$.

Under the above assumptions, we have the following results.

**Lemma 2.6.** If Assumption 2.3 holds, then for any $x \in \mathcal{X}$, $\hat{p}(x,t)$ is non-decreasing in $t$ when $t > 0$.

**Lemma 2.7.** Suppose that Assumptions 2.1 to 2.3 are satisfied. Then, for any $t \in (-\delta, \delta)$ and $x_0 \in \mathcal{O}$,

(i): \[
\partial_x g_1(x_0, t) = \mathbb{E}[\text{conv} \{\nabla_x c_i(x_0, \xi) : i \in I(x_0, t, \xi)\}],
\]

where \[I(x_0, t, \xi) = \{i : c_i(x_0, \xi) + t = \phi(c(x_0, \xi), t)\}.
\]

(ii): $g_1(x, \cdot)$ is continuously differentiable in $t$, and

$$\frac{\partial}{\partial t} g_1(x, t_0) = 1 - F(-t_0, x).$$

**Proof.** Note $g_1(x, t)$ is convex in $(x, t)$ and $x_0 \in \text{int} (\text{dom } g_1(\cdot, t))$, then by [12, Theorem 7.47] and the formula on subgradient of convex function in [10], the result in (i) is easy to be obtained.

Let $\psi(t) = g_1(x, t)$. Then by [12, Theorem 7.44], $\psi$ is directional differentiable at $t_0$,

$$\psi'(t_0; 1) =\mathbb{E}[1_{\{c(x, \xi) + t_0 \geq 0\}}] = \Pr\{c(x, \xi) + t_0 \geq 0\} = (1 - F(-t_0, x)) \times 1.$$  

Similarly,

$$\psi'(t_0; -1) = (1 - F(-t_0, x)) \times (-1).$$

Since $F(\cdot, \cdot)$ is continuous, we have $\psi(\cdot)$ is differentiable at $t_0$, and $\psi'(t_0) = 1 - F(-t_0, x)$. The proof is complete. \qed
2.2. ε-approximation. The following theorem is critical which is the base of ε-approximation theory.

**Theorem 2.8.** Suppose that Assumptions 2.1 to 2.3 are satisfied. Then,

\[
p(x) = \inf_{t > 0} \hat{p}(x, t) = \lim_{t \downarrow 0} \hat{p}(x, t).
\]

**Proof.** Since \( g_1(x, t) \) is convex and directional differentiable, then

\[
\lim_{t \downarrow 0} \frac{g_1(x, t) - g_1(x, 0)}{t} = \lim_{t \downarrow 0} \hat{p}(x, t).
\]

Then by Lemma 2.6 and 2.7(ii),

\[
\partial \hat{p}(x, 0) = \inf_{t > 0} \hat{p}(x, t) = 1 - F(0, x) = p(x).
\]

The proof is complete. \( \square \)

By Theorem 2.8, for \( \varepsilon > 0 \) small enough, \( \hat{p}(x, \varepsilon) \leq 1 - \alpha \) is a conservative approximation of \( p(x) \leq 1 - \alpha \). Since \( \hat{p}(x, \varepsilon) \leq 1 - \alpha \) is equivalent to

\[
g_1(x, \varepsilon) - g_2(x) \leq \varepsilon(1 - \alpha),
\]

the approximation problem based on the constraint, denoted by \( (P_\varepsilon) \), is in the form of

\[
\begin{align*}
\text{minimize} & \quad f_1(x) - f_2(x) \\
\text{subject to} & \quad g_1(x, \varepsilon) - g_2(x) \leq \varepsilon(1 - \alpha), \\
& \quad x \in \mathcal{X}.
\end{align*}
\]

In the rest of this section, we show that, when \( \varepsilon > 0 \) is small enough, the DC optimization Problem \( (P_\varepsilon) \) is a good approximation of Problem \( (P) \).

Let \( \Omega(\varepsilon) = \{x \in \mathcal{X} : g_1(x, \varepsilon) - g_2(x) \leq \varepsilon(1 - \alpha)\} \) represents the feasible set of Problem \( (P_\varepsilon) \). For sets \( A, B \), let

\[
dist(x, A) := \inf_{y \in A} \|x - y\|
\]
denote the distance from \( x \) to \( A \), and

\[
\mathbb{D}(A, B) := \sup_{x \in A} \dist(x, B)
\]
denote the deviation of set \( A \) from set \( B \). Let \( S(\varepsilon) \) and \( \nu(\varepsilon) \) be the set of optimal solutions and the optimal value of Problem \( (P_\varepsilon) \), \( S_0 \) and \( \nu_0 \) be the set of optimal solutions and the optimal value of Problem \( (P) \). Then we have the following results.

**Theorem 2.9.** [6, Lemma 3, Theorem 2] Suppose that Assumptions 2.1 to 2.4 are satisfied. Then, \( \lim_{\varepsilon \downarrow 0} \Omega(\varepsilon) = \Omega_0 \), \( \lim_{\varepsilon \downarrow 0} \mathbb{D}(S(\varepsilon), S_0) = 0 \) and \( \lim_{\varepsilon \downarrow 0} \nu(\varepsilon) = \nu_0 \).

Theorem 2.9 shows that when \( \varepsilon > 0 \) is small enough, the optimal solution of Problem \( (P_\varepsilon) \) is a good approximation to Problem \( (P) \). However, since Problem \( P_\varepsilon \) is a DC program, we can only obtain a stationary point. Thus, we should discuss whether the stationary point of Problem \( (P_\varepsilon) \) converges to the stationary point of \( (P) \) when \( \varepsilon \rightarrow 0 \).

Under Assumption 2.4, i.e., \( g_2(\cdot) \) is differentiable on \( \mathcal{O} \), define

\[
\partial p(x) = \lim_{y \rightarrow x} \sup_{t \downarrow 0} \frac{\partial \hat{p}(y, t) - \nabla g_2(y)}{t}.
\]
Let \( \Lambda_0 \) and \( \Lambda_\varepsilon \) be the sets of stationary points of Problem (P) and (P_\varepsilon), respectively, namely,

\[
\Lambda_0 = \{(x, \lambda) \in \Omega_0 \times \mathbb{R}_+ : 0 \in \partial f(x) + \lambda \partial p(x) + N_X(x), \lambda(p(x) - 1 + \alpha) = 0 \}
\]

and

\[
\Lambda_\varepsilon = \{(x, \lambda) \in \Omega_0 \times \mathbb{R}_+ : 0 \in \partial f(x) + \lambda \partial p(x, \varepsilon) + N_X(x), \lambda(p(x, \varepsilon) - 1 + \alpha) = 0 \}
\]

**Theorem 2.10.** Suppose that Assumptions 2.1 to 2.5 are satisfied. Then

\[
\lim_{\varepsilon \downarrow 0} D(\Lambda_\varepsilon, \Lambda_0) = 0
\]

**Proof.** For any \((x, \lambda) \in \limsup_{\varepsilon \downarrow 0} \Lambda_{\varepsilon_k} \), there exists \((x_k, \lambda_k) \in \Lambda_{\varepsilon_k} \) such that \((x_k, \lambda_k) \to (x, \lambda) \).

According to the outer semicontinuity of \(N_X\) and \(\partial f\), and the definition of \(\Lambda_{\varepsilon_k}\), take \(k \to \infty\), we have

\[
0 \in \lim_{k \to \infty} \left[ \partial f(x_k) + \lambda_k \frac{\partial g_1(x_k, \varepsilon_k) - \nabla g_2(x_k)}{\varepsilon_k} \right] + N_X(x_k)
\]

\[
\subset \partial f(x) + \lambda \partial p(x) + N_X(x). \tag{4}
\]

By Lemma 2.7 (ii), for any \(x' \in \mathcal{O}, t \to g_1(x', t)\) is continuously differentiable, there exist \(\tilde{\varepsilon}_k \in (0, \varepsilon_k), k = 1, 2, \ldots\), such that

\[
\frac{g_1(x_k, \varepsilon_k) - g_2(x_k)}{\varepsilon_k} = \frac{\partial}{\partial t} g_1(x_k, \tilde{\varepsilon}_k) = 1 - F(-\tilde{\varepsilon}_k, x_k).
\]

Under Assumption 2.3, \(F(t, x)\) is continuous in \((-\delta, \delta) \times \mathcal{O}\), then

\[
\lim_{k \to +\infty} \frac{g_1(x_k, \varepsilon_k) - g_2(x_k)}{\varepsilon_k} = \lim_{k \to +\infty} [1 - F(-\tilde{\varepsilon}_k, x_k)] = 1 - F(0, x) = p(x). \tag{5}
\]

By the definition of \(\Lambda_0\) we get \((x, \lambda) \in \Lambda_0\). Thus \(\limsup_{\varepsilon \downarrow 0} \Lambda(\varepsilon) \subset \Lambda_0\). The proof is complete. \(\square\)

3. Sequential convex program method for problem (P_\varepsilon). Problem (P_\varepsilon) is a DC program, because the object function \(f_1(x) - f_2(x)\) and the left hand side of the constraint \(g_1(x, \varepsilon) - g_2(x) \leq \varepsilon(1 - \alpha)\) are both DC functions. In this section, we propose a sequential convex program method to obtain a stationary point of Problem (P_\varepsilon).

Define

\[
g(x, \varepsilon) = g_1(x, \varepsilon) - g_2(x) - \varepsilon(1 - \alpha)
\]

and

\[
\Omega(\varepsilon) = \{x \in \mathcal{X} : g(x, \varepsilon) \leq 0\}.
\]

For \(y \in \mathcal{X}\), define

\[
\Omega(\varepsilon)_{y, w} = \{x \in \mathcal{X} : g_1(x, \varepsilon) - [g_2(y) + \langle w, x - y \rangle] - \varepsilon(1 - \alpha) \leq 0\},
\]

where \(w \in \partial g_2(y)\). Since \(g_2(x)\) is convex, for any \(y \in \mathcal{X}\) and \(w \in \partial g_2(y)\), we have

\[
g_1(x, \varepsilon) - g_2(x) \leq g_1(x, \varepsilon) - [g_2(y) + \langle w, x - y \rangle].
\]

Then

\[
\Omega(\varepsilon)_{y, w} \subset \Omega(\varepsilon). \tag{6}
\]

For \(y \in \mathcal{X}\), \(\Omega(\varepsilon)_{y, w}\) is a convex subset of \(\Omega(\varepsilon)\), since \(g_1(x, \varepsilon) - [g_2(y) + \langle w, x - y \rangle] - \varepsilon(1 - \alpha)\) is convex in \(x\). Define Problem CP(\varepsilon, y, w, v) as
Assume that for any Assumption 3.2.

Then for \( y \in X \), Problem \( \text{CP}(\varepsilon, y, w, v) \) is a conservative convex approximation of Problem \((P_\varepsilon)\). By solving a sequence of Problem \( \text{CP}(\varepsilon, y, w, v) \), we can obtain a stationary point of Problem \((P_\varepsilon)\).

**Algorithm 1 (Sequential Convex Approximation Method)**

**Step 0:** Give \( x^0 \in \Omega(\varepsilon) \). Let \( k = 0 \).

**Step 1:** Stop if

\[
0 \in \partial f(x^k) + N_{\mathbb{R}_-}(g(x^k, \varepsilon)) \overline{\partial} g(x^k, \varepsilon) + N_X(x^k),
\]

where \( N_{\mathbb{R}_-}(g(x^k, \varepsilon)) \) is the norm cone of \( \mathbb{R}_- \) at \( g(x^k, \varepsilon) \), and

\[
N_{\mathbb{R}_-}(g(x^k, \varepsilon)) \overline{\partial} g(x^k, \varepsilon) = \{ \lambda v : v \in \overline{\partial} g(x^k, \varepsilon), 0 \leq \lambda \perp g(x^k, \varepsilon) \leq 0 \}.
\]

**Step 2:** Choose \( v^k \in \partial f_2(x^k), w^k \in \partial g_2(x^k) \), solve Problem \( \text{CP}(\varepsilon, x^k, w^k, v^k) \) to obtain its optimal solution \( x^{k+1} \).

**Step 3:** Set \( k = k + 1 \), go to Step 1.

**Remark 1.** For Algorithm 1, the following works in the literature should be mentioned.

- For the case \( f_2 = 0 \), i.e., the object function is convex, Hong et al. [6] provided a comprehensive analysis on the convergence of the algorithm.
- In [3], a low rank constrained optimization problem is discussed, and the technique by linearizing the second convex term of DC function also plays an important role in the algorithm design.

**Lemma 3.1.** If \( x^0 \in \Omega(\varepsilon) \) and \( \{ x^k \} \) is generated by Algorithm 1 for Problem \((P_\varepsilon)\), then \( \{ x^k \} \subset \Omega(\varepsilon) \), and \( \{ f(x^k) \} \) is a convergent non-increasing sequence.

**Proof.** Note that, by (6), \( x^1 \in \Omega(\varepsilon) \subset \Omega(\varepsilon) \). Therefore, by the principle of induction, \( \{ x^k \} \subset \Omega(\varepsilon) \).

As \( f(x) \) is continuous and \( X \) is compact, we obtain \( \inf_{x \in X} h(x) \) is finite. Since \( \{ x^k \} \subset \Omega(\varepsilon) \subset X \), for all \( k \), we have

\[
f(x^k) \geq \inf_{x \in \Omega(\varepsilon)} f(x),
\]

which implies \( \{ f(x^k) \} \) is lower bounded. For any \( k \geq 1 \), \( x^{k+1} \in \text{argmin}\{ f_1(x) - (f_2(x) + \langle v^k, x - x^k \rangle) : x \in \Omega(\varepsilon) \} \), we have

\[
f(x^{k+1}) = f_1(x^{k+1}) - f_2(x^{k+1}) \\
\leq f_1(x^{k+1}) - [f_2(x^k) + \langle v^k, x^{k+1} - x^k \rangle] \\
\leq f_1(x^k) - [f_2(x^k) + \langle v^k, x^k - x^k \rangle] \\
= f(x^k),
\]

Then \( \{ f(x^k) \} \) is a non-increasing lower bounded set, which is convergent. \( \square \)

To get the convergence result of the above algorithm, we need the following assumption.

**Assumption 3.2.** Assume that for any \( x \in \Omega_0 \),

\[
0 \notin \limsup_{t \downarrow 0, x' \to x} \partial x g_1(x', t) - \partial g_2(x) \parallel N_X(x). \tag{7}
\]
Lemma 3.3. Suppose that Assumptions 2.1 to 2.4 are satisfied, and Assumption 3.2 holds. If \( g_2 \) is strictly convex on \( \mathcal{X} \), then for \( \varepsilon > 0 \) small enough, \( y \in \Omega(\varepsilon) \) and \( w \in \partial g_2(y) \), there exists \( \gamma > 0 \) such that
\[
\inf_{x \in \text{bd} \Omega(\varepsilon)_{y,w}} \min_{u \in \partial \gamma_1(\varepsilon,y,w,x)} \| u \| \geq \gamma,
\]
where
\[
\gamma_1(\varepsilon, y, w, x) = g_1(x, \varepsilon) - [g_2(y) + \langle w, x - y \rangle] - \varepsilon(1 - \alpha) + \delta_\mathcal{X}(x).
\]

Proof. Suppose there exist \( \varepsilon_i \downarrow 0 \), \( y_i \in \Omega(\varepsilon_i) \), \( w^i \in \partial g_2(y^i) \), \( x^i \in \text{bd} \Omega(\varepsilon_i)_{y^i,w^i} \), and \( u^i \in \partial \gamma_1(\varepsilon_i, y^i, w^i, x^i) \), such that \( u^i \to 0 \). Since \( X \) is compact, we assume that \( y^i \to \hat{y}, \) \( w^i \to \hat{w} \) and \( x^i \to \hat{x} \). Next we prove \( \hat{y} = \hat{x} \).

Because \( x^i \in \text{bd} \Omega(\varepsilon_i)_{y^i,w^i} \), we can assume that
\[
g_1(x^i, \varepsilon_i) - [g_2(y^i) + \langle w^i, x^i - y^i \rangle] = \varepsilon_i(1 - \alpha), \quad x^i \in \mathcal{X}.
\]
Let \( i \to \infty \), we have
\[
g_1(\hat{x}, 0) - g_2(\hat{y}) + \langle \hat{w}, \hat{x} - \hat{y} \rangle = 0.
\]
Note \( g_2 \) is strictly convex, then \( \hat{x} = \hat{y} \). There exist \( \xi^i \in \partial g_1(x^i, \varepsilon_i) \) and \( \eta^i \in N_X(x^i) \) such that
\[
u^i = \xi^i - \nu^i + \eta^i.
\]
Assume that \( \xi^i \to \hat{\xi}, \) \( \eta^i \to \hat{\eta}, \) then
\[
\hat{\xi} \in \lim_{\xi \to 0} \partial g_1(x^i, t), \quad \hat{\eta} \in N_X(\hat{x}), \quad \hat{w} \in \partial g_2(\hat{x}).
\]
By (9), we have \( \hat{\xi} - \hat{\eta} = -\hat{\eta} \in -N_X(\hat{x}) \), which is in contradiction with Assumption 3.2.

The condition (8) is called asymptotic constraint qualification, which is a stronger constraint qualification than Slater condition, see \cite[Theorem 2]{8}. By \cite[Theorem 4]{8}, if the condition (8) is satisfied, the following metric regularity holds,
\[
dist(x, \Omega(\varepsilon)_{y,w}) \leq \frac{1}{\gamma}[g_1(x) - (g_2(y) + \langle v, x - y \rangle)] + \delta_\mathcal{X}(x), \quad \forall x \in \mathbb{R}^n.
\]

Lemma 3.4. For Problem \((P_\varepsilon)\), assume that \( \{y^k\} \subset \mathcal{X} \) is a sequence converges to \( \hat{y} \in \Omega(\varepsilon) \), \( w^k \in \partial g_2(y^k) \) and \( w \in \partial g_2(\hat{y}) \) satisfied \( w^k \to w \). If \( \Omega(\varepsilon)_{\hat{y},w} \) is a nonempty compact set, and condition (8) holds, then \( \lim_{k \to +\infty} \Omega(\varepsilon)_{y^k,w^k} = \Omega(\varepsilon)_{\hat{y},w} \).

Proof. Note that
\[
\liminf_{k \to +\infty} \Omega(\varepsilon)_{y^k,w^k} \subset \limsup_{k \to +\infty} \Omega(\varepsilon)_{y^k,w^k},
\]
we only need to prove
\[
\limsup_{k \to +\infty} \Omega(\varepsilon)_{y^k,w^k} \subset \Omega(\varepsilon)_{\hat{y},w} \subset \liminf_{k \to +\infty} \Omega(\varepsilon)_{y^k,w^k}.
\]
Let \( z \in \limsup_{k \to +\infty} \Omega(\varepsilon)_{y^k,w^k} \), then there exist a sequence \( \{k_j\} \subset \mathbb{N} \) and \( z_j \in \Omega(\varepsilon)_{y^{k_j},w^{k_j}} \) satisfying
\[
z = \lim_{j \to \infty} z_j, \quad g_1(z_j) - [g_2(y^{k_j}) + \langle w^{k_j}, z_j - y^{k_j} \rangle] \leq \varepsilon(1 - \alpha).
\]
Let \( j \to +\infty \), we obtain \( z \in \mathcal{X} \), and
\[
g_1(z) - \left[ g_2(\bar{y}) + \langle w, z - \bar{y} \rangle \right] \leq \epsilon(1 - \alpha),
\]
thus \( z \in \Omega(\epsilon_{\bar{y}, w}) \). Then the result \( \limsup_{k \to +\infty} \Phi_{\epsilon_k}(V^k) \subset \Omega(\epsilon_{\bar{y}, w}) \) is proved.

In the following we will prove \( \Omega(\epsilon_{\bar{y}, w}) \subset \liminf_{k \to +\infty} \Omega(\epsilon_{\bar{y}, w^k}) \). Let \( z \in \Omega(\epsilon_{\bar{y}, w}) \), we only need to show there exists \( \{z_j\} \) such that, for \( j \) large enough, \( z_j \in \Omega(\epsilon_{\bar{y}, w^j}) \) and \( z_j \to z \).

Let
\[
\psi(x) = g_1(x) - \left[ g_2(\bar{y}) + \langle w, x - \bar{y} \rangle \right],
\]
\[
\psi^k(x) = g_1(x) - \left[ g_2(y^k) + \langle w^k, x - y^k \rangle \right].
\]
Define \( d^k(x) = \psi^k(x) - \psi(x) \), then
\[
\|d^k(y') - d^k(y'')\| = \|(u^k - w)(y' - y'')\| \leq \|w^k - w\| \|y' - y''\|, \quad \forall y', y'' \in \mathcal{X}.
\]
Since \( w^k \) converges to \( w \), let \( k \) large enough such that
\[
\gamma : = \|w^k - w\| \leq \frac{\gamma}{4}.
\]
Then for any \( x \in \mathcal{X} \), we have \( \partial d^k(x) \subset \frac{\gamma}{4} \mathcal{B} \), which implies that
\[
\partial \psi_x^k \subset \partial \psi(x) + \frac{\gamma}{4} \mathcal{B}.
\]
(11)
The condition (8) holds for \( \Omega(\epsilon_{\bar{y}, w}) \), then
\[
\|u\| \geq \gamma, \quad \forall u \in \partial \psi(x), \forall x \in \partial \Omega(\epsilon_{\bar{y}, w}).
\]
According to the outer semicontinuity of \( \partial \psi(\cdot) \), for any \( x \in \partial \Omega(\epsilon_{\bar{y}, w}) \), there exists \( \delta_x > 0 \) such that
\[
\|u\| \geq \frac{\gamma}{2}, \quad \forall u \in \partial \psi(x'), \forall x' \in \mathcal{B}_{\delta_x}(x).
\]
(12)
(If \( x' \notin X \), then \( \partial \psi(x') = \emptyset \), (12) always holds.) Because \( \partial \Omega(\epsilon_{\bar{y}, w}) \) is compact, \( \{\mathcal{B}_{\delta_x/2}(x) : x \in \partial \Omega(\epsilon_{\bar{y}, w})\} \) is an open cover of \( \partial \Omega(\epsilon_{\bar{y}, w}) \), based on the finite set covering theorem, there exist finite points in \( \partial \Omega(\epsilon_{\bar{y}, w}) \), for example \( x^1, \ldots, x^K \), such that
\[
\partial \Omega(\epsilon_{\bar{y}, w}) \subset \bigcup_{i=1}^K \mathcal{B}_{\delta_{x^i}/2}(x^i).
\]
Let
\[
\delta = \min_{1 \leq i \leq K} \frac{\delta_{x^i}}{4},
\]
then for \( k \) large enough,
\[
\partial \Omega(\epsilon_{\bar{y}, w^k}) \subset \partial \Omega(\epsilon_{\bar{y}, w}) + \delta \mathcal{B},
\]
and
\[
\|u\| \geq \frac{\gamma}{2}, \quad \forall u \in \partial \psi(x'), \forall x' \in \partial \Omega(\epsilon_{\bar{y}, w^k}).
\]
(13)
Together with (11) and (13), for \( k \) large enough, we have
\[
\inf_{x \in \partial \Omega(\epsilon_{\bar{y}, w^k})} \min_{u \in \partial \psi^k(x)} \|u\| \geq \frac{\gamma}{4}.
\]
(14)
By (10), for \( k \) large enough,
\[
\text{dist}(x, \Omega(\epsilon_{\bar{y}, w^k})) \leq \frac{4}{\gamma} \|\psi^k(x)\|, \forall x \in \mathbb{R}^n.
\]
(15)
Then for $k$ large enough, there exists $z^k \in \Omega(\varepsilon)_{g^k, w^k}$ such that
\[
\|z^k - z\| \leq \frac{4}{\gamma} \langle \psi^k(z) \rangle + \frac{4}{\gamma} \langle \psi(z) + d^k(z) \rangle.
\] (16)

Note, for $k$ large enough such that $\|w^k\| \leq 2\|w\|$, we obtain
\[
\|d^k(z)\| = \|g_2(\bar{y}) - g_2(y^k) + \langle w, z - \bar{y} \rangle - \langle w^k, z - y^k \rangle\|
\leq \|g_2(\bar{y}) - g_2(y^k)\| + \|w - w^k\| \|z\| + \|w\| \|y^k - \bar{y}\| + \|w^k\| \|y^k - \bar{y}\|,
\]
where $M_0 = \max\{\|y\| : y \in \Omega(\varepsilon)_{\bar{g}, w}\}$, by (16) we have
\[
\|z^k - z\| \leq \frac{4}{\gamma} \langle \psi(z) \rangle + 4 \|g_2(\bar{y}) - g_2(y^k)\| + M_0 \|w - w^k\| + 3 \|w\| \|y^k - \bar{y}\|,
\]
which implies $z^k \to z$. Thus $\Omega(\varepsilon)_{\bar{g}, w} \subset \inf_{k \to +\infty} \Omega(\varepsilon)_{g^k, w^k}$. The proof is complete.

Then we have the following theorem on the convergence of the sequential convex program method.

**Theorem 3.5.** Assume the sequence $\{x^k\}$ is generated by Algorithm 1 to solve Problem ($P_\varepsilon$). Suppose Assumptions 2.1 to 2.4 are satisfied, and Assumption 3.2 holds. If $g_2$ is strictly convex on $X$, then for $\varepsilon > 0$ small enough, we have $\{x^k\} \subset \Omega(\varepsilon)$, $\{f(x^k)\}$ is non-decreasing, every cluster point of $\{x^k\}$ is a stationary point of Problem ($P_\varepsilon$).

**Proof.** Suppose that $\bar{x}, w, v$ is a cluster point of the sequence $\{x^k, w^k, v^k\}$. By Lemma 3.1, we obtain $\{x^k\} \subset \Omega(\varepsilon)$, then the cluster point of $\{x^k\}$ satisfies $\bar{x} \in \Omega(\varepsilon)$. Since $\partial f_2$ and $\partial g_2$ is outer semicontinuous at $\bar{x}$, there exists $\delta_1 > 0$, such that for $j \in \mathbb{N}$ large enough, $x^j \in \{x^k\} \cap B_{\delta_1}(\bar{x})$, $\partial f_2(x^j)$ and $\partial h(x^j)$ are bounded, $v \in \partial f_2(\bar{x})$ and $w \in \partial g_2(\bar{x})$.

By Lemma 3.1, $\{f(x^k)\}$ is monotone non-increasing, which converges to $\inf_{x} f(x^k)$, then $f(\bar{x}) = \inf_{x} f(x^k)$. Since the Assumption 3.2 is satisfied, the condition (8) holds for $\Omega(\varepsilon)_{\bar{g}, w}$. We first prove $\bar{x}$ is an optimal point of Problem CP$(\varepsilon, \bar{x}, w, v)$. Let $\{k_j\}$ is an infinite sequence of $\mathbb{N}$ such that $(x^{k_j}, w^{k_j}, v^{k_j}) \to (\bar{x}, w, v)$. Define
\[
\bar{f}(x) = f_1(x) - \|f_2(\bar{x}) + \langle v, x - \bar{x} \rangle\| + \delta_{\Omega(\varepsilon)_{\bar{x}, w}}(x),
\]
and
\[
\bar{f}_{k_j}(x) = f_1(x) - \|f_2(x^{k_j}) + \langle v^{k_j}, x - x^{k_j} \rangle\| + \delta_{\Omega(\varepsilon)_{x^{k_j}, w^{k_j}}}(x).
\]

By Lemma 3.4 we have
\[
\Omega(\varepsilon)_{x^{k_j}, w^{k_j}} \to \Omega(\varepsilon)_{\bar{x}, w}.
\]
By [11, Proposition 7.4], $\delta_{\Omega(\varepsilon)_{x^{k_j}, w^{k_j}}}(\cdot)$ epi-converges to $\delta_{\Omega(\varepsilon)_{\bar{x}, w}}(\cdot)$, then we get $\bar{f}_{k_j}(\cdot)$ epi-converges $\bar{f}(\cdot)$. By [11, Theorem 7.33], we obtain
\[
\inf_{x} \bar{f}_{k_j}(x) \to \inf_{x} \bar{f}(x),
\]
and
\[
\lim_{j \to +\infty} \sup_{x} \arg\min \bar{f}_{k_j} \subset \arg\min \bar{f}.
\] (17)

Note that $\arg\min \bar{f}_{k_j}$ is the optimal set of Problem CP$(\varepsilon, x^{k_j}, w^{k_j}, v^{k_j})$, $\arg\min \bar{f}$ is the optimal set of Problem CP$(\varepsilon, \bar{x}, w, v)$. By (17), the sequence $\{x^{k_j+1}\}$ has a cluster point $\tilde{x} \in \Omega(\varepsilon)$, which is in the set $\arg\min \bar{f}$. Since $f(\tilde{x}) \leq f(\bar{x})$ and
We consider the problem

$$\minimize x \in X \quad f(x) = \inf_k f(x^k),$$

we have $f(\bar{x}) = f(\hat{x})$. Then $\bar{x}$ is the optimal solution of Problem CP($\varepsilon, \bar{x}, w, v$). Because Problem CP($\varepsilon, \bar{x}, w, v$) is a convex problem and Slater condition holds, we have

$$0 \in \partial f(\bar{x}) - v + [\partial g(\bar{x}) - w]N_{\mathbb{R}^p}(\bar{g}(\bar{x})) + N_{X}(\bar{x}),$$

which implies $\bar{x}$ is a stationary point of Problem (P$\varepsilon$).

4. Numerical experiments. In this section, we will first consider some numerical issues for implementing our algorithm and then show the results for some numerical examples.

4.1. Monte Carlo method. To implement Algorithm 1, at each iteration, we need to solve a convex program

$$\begin{align*}
\minimize f_1(x) - (f_2(y) + \langle v, x - y \rangle) \\
\text{subject to} \quad g_1(x, \varepsilon) - [g_2(y) + \langle w, x - y \rangle] &\leq \varepsilon(1 - \alpha) \\
x &\in X.
\end{align*}$$

(19)

The problem is difficult to solve because $g_1(x, \varepsilon)$ and $g_2(x)$ are expressed by mathematical expectations, and the closed forms of them are unknown. To overcome this difficulty, we apply a Monte Carlo method.

Recall the definitions of $g_1, g_2$ and Lemma 2.7, we have

$$g_1(x, \varepsilon) = \mathbb{E}[\varepsilon + c(x, \xi)]_+,$$

$$g_2(y) = \mathbb{E}[c(y, \xi)]_+,$$

$$w \in \partial g_2(y) = \mathbb{E}[\text{conv}\{\nabla_x c_i(y, \xi) : i \in I(y, \xi)\}],$$

where $c(x, \xi) = \max\{c_i(x, \xi)\}$ and $I(y, \xi) = \{i : c_i(y, \xi) = [c(y, \xi)]_+\}$. Let $w(y, t)$ be a point of the convex set $\text{conv}\{\nabla_x c_i(y, t) : i \in I(y, t)\}$. Let

$$g(x) = g_1(x, \varepsilon) - [g_2(y) + \langle w, (x - y) \rangle].$$

Let $\xi_1, \ldots, \xi_n$ denote an independent and identically distributed (i.i.d.) sample of $\xi$. Then, a natural estimator of $g(x)$ is

$$\hat{g}(x) = \frac{1}{n} \sum_{k=1}^{n} \{c(x, \xi_k) + \varepsilon - [c(y, \xi_k)]_+ - \langle w(y, \xi_k), x - y \rangle\}.$$

We consider the problem

$$\begin{align*}
\minimize \quad f_1(x) - (f_2(y) + \langle v, x - y \rangle) \\
\text{subject to} \quad \hat{g}(x) &\leq \varepsilon(1 - \alpha) \\
x &\in X,
\end{align*}$$

(20)

and use its optimal solution to approximate the optimal solution of Problem (19). Let $S$ and $\nu^*$ denote the set of optimal solutions and the optimal objective value of Problem (19), and $\hat{S}_n$ and $\hat{\nu}_n^*$ denote the set of optimal solutions and the optimal objective value of Problem (20). Then we have the following lemma to show that $\hat{S}_n$ and $\hat{\nu}_n^*$ converge to $S$ and $\nu^*$.

Lemma 4.1. [6, Theorem 6] Suppose that Assumptions 2.1 to 2.5 are satisfied. When $\varepsilon > 0$ is small enough, $\mathbb{D}(\hat{S}_n, S) \rightarrow 0$ w.p.1 and $\hat{\nu}_n^* \rightarrow \nu^*$ w.p.1 as $n \rightarrow \infty$.

4.2. Numerical environment. The numerical experiment is carried out in MATLAB 7.10 running on a PC Intel Core i5 CPU (3.20 GHz) and 4 GB RAM. In the following test problems, we set $\alpha = 0.9$, $\varepsilon = 0.05^2$. Let $\{x^k\}$ be the generated sequence. We stop the algorithm if norm($x^{k+1} - x^k$)$^2 \leq 10^{-6}$.
4.3. A $L_1$ optimization problem. We consider the following problem:

\[
\begin{align*}
\text{minimize} \quad & -\|x\|_1 \\
\text{subject to} \quad & \Pr\{\sum_{j=1}^n \xi_{ij}^2 x_j^2 - b \leq 0, \ i = 1, \ldots, m\} \leq \alpha,
\end{align*}
\]

where $\xi_{ij}, i = 1, \ldots, m$ and $j = 1, \ldots, n$ are independent and identically distributed standard norm random variables, $b$ is a constant. The Problem (21) can be written equivalently as

\[
\begin{align*}
\text{minimize} \quad & -\sum_{j=1}^n x_j \\
\text{subject to} \quad & \Pr\{\sum_{j=1}^n \xi_{ij}^2 x_j^2 - b \leq 0, \ i = 1, \ldots, m\} \leq \alpha, \quad x \geq 0,
\end{align*}
\]

in the sense that Problem (21) and Problem (22) have the same optimal value, and the optimal set for Problem (22) is a subset of that for Problem (21). Note that Problem (22) is a simple type of JCCDCP (the object function is linear).

**Lemma 4.2.** The optimal solution $x^*$ of Problem (22) is

\[
x_1^* = x_2^* = \cdots = x_n^* = \left[\frac{b}{F_{\chi_n^2}^{-1}(\alpha \frac{n}{m})}\right]^\frac{1}{2},
\]

where $F_{\chi_n^2}^{-1}$ is the inverse chi-squared distribution function with $n$ degrees of freedom.

**Proof.** By the symmetric form of Problem (22), the optimal solution $x^*$ satisfies $x_1^* = x_2^* = \cdots = x_n^*$. Then the joint chance constraint of Problem (22) in $x^*$ takes the form

\[
\left[\Pr\left\{\left(x_1^*\right)^2 \sum_{j=1}^n \xi_{ij}^2 \leq b \right\}\right]^m \leq \alpha,
\]

which is equivalent to

\[
F_{\chi_n^2}\left[\frac{b}{(x_1^*)^2}\right] \geq \alpha^{\frac{1}{m}}.
\]

Then, it is clear that

\[
x_1^* = \left[\frac{b}{F_{\chi_n^2}^{-1}(\alpha \frac{n}{m})}\right]^\frac{1}{2}.
\]

\[\square\]

We implement Algorithm 3 for the following two problems. At each iteration in the algorithm, we use CVX [5] to solve a convex program generated by the Monte Carlo method.

**Problem 4.3.** In Problem (22), let $n = 3, m = 2, b = 9, N = 100$. By Lemma 4.2, the optimal solution is $(1.08, 1.08, 1.08)^T$, the optimal value is -3.23.

**Problem 4.4.** In Problem (22), let $n = 10, m = 10, b = 100, N = 100$. By Lemma 4.2, the optimal solution is $(2.08, 2.08, \cdots, 2.08)^T$, the optimal value is -20.82.

The numerical results are shown in Figure 2. Take Problem 4.4 for example, the algorithm typically requires less than 20 iterations to converge to the optimal value. When sample size $N = 100$, at each iteration, a convex program with 4223 variables and 3112 constraints is solved in 26 seconds on average. When $N$ grows, the size of the convex program increases much faster, and we found CVX performed not well under this situation.
4.4. A nonconvex quadratic programming problem. Let \( G \in \mathbb{R}^{n \times n} \) be a symmetric but not positive semidefinite matrix. We consider the following nonconvex stochastic quadratic programming problem,

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} x^T G x + c^T x \\
\text{subject to} & \quad \Pr\{\Xi x \leq b\} \leq \alpha \\
& \quad l \leq x \leq u,
\end{align*}
\]

where

\[
\Xi = \begin{bmatrix}
\xi_1 \\
\xi_2 \\
\vdots \\
\xi_m
\end{bmatrix}
\]

with \( \xi_i = (\xi_{i1}, \ldots, \xi_{in}) \) be a \( m \times n \) matrix of random variables, \( c, l, u \in \mathbb{R}^n \) and \( b \in \mathbb{R}^m \). Assume that \( \lambda \) is the largest eigenvalue of \( G \). Let \( \rho > \max(0, \lambda) \) be a positive number and \( I \in \mathbb{R}^{n \times n} \) be the identity matrix. Define

\[
f_1(x) := \frac{1}{2} \rho \| x \|^2 + c^T x
\]

and

\[
f_2(x) := \frac{1}{2} x^T (\rho I - G) x.
\]

Note that both \( f_1 \) and \( f_2 \) are convex. Then the Problem (24) can be reformulate as

\[
\begin{align*}
\text{minimize} & \quad f_1(x) - f_2(x) \\
\text{subject to} & \quad \Pr\{\xi_i x - b_i \leq 0, \ i = 1, \ldots, m\} \leq \alpha, \\
& \quad l \leq x \leq u,
\end{align*}
\]

which is a JCCDCP.

**Problem 4.5.** In Problem (24), let \( \xi_{ij}, i = 1, \ldots, m \) and \( j = 1, \ldots, n \) be independent and identically distributed standard norm random variables, and \( G \) be randomly generated by MATLAB.

The numerical results for Problem 4.5 are reported in Table 1, where “It.”, “Res0.”, and “Res*.” stand for the number of iterations, the residuals \( \phi(\cdot, \cdot, \cdot) \) at the first iteration and at the final iteration, respectively.

For Problem 4.5, we use \texttt{fmincon}, which is a function from MATLAB Optimization Toolbox, to solve the convex program at each iteration in the algorithm. Problem
Table 1. Numerical results for Problem 4.5

| N  | m  | n  | cputime | It. | Res0. | Res*  |
|----|----|----|---------|-----|-------|-------|
| 100| 5  | 20 | 0:09    | 9   | 0.8e-1| 0.9e-5|
| 20 | 100| 15 | 1.29    | 15  | 1.0e-1| 0.7e-5|
| 40 | 200| 15 | 5.00    | 15  | 0.9e-1| 0.8e-5|
| 50 | 500| 15 | 23.55   | 15  | 0.8e-1| 0.8e-5|
| 1000| 5 | 20 | 1:46    | 13  | 1.2e-1| 0.9e-5|
| 20 | 100| 16 | 12.52   | 16  | 1.0e-1| 0.9e-5|
| 40 | 200| 14 | 43.55   | 14  | 0.9e-1| 0.8e-5|
| 10000| 5 | 20 | 23:04   | 22  | 2.4e-1| 0.8e-5|
| 20 | 100| 16 | 4:21:41 | 16  | 1.0e-1| 0.9e-5|

4.5 is a nonconvex stochastic program and the optimal solution is no longer known. However, from the numerical results in Table 1, by our method we can always obtain a stationary point in a suitable time even for a medium scale problem. Also from Table 1, as the sample size $N$ grows, the size of the convex program in each iteration increases fast, consequently the CPU time becomes much longer.

5. Conclusions. In this paper, we introduced a DC technique to approximate the DC programming problem with joint chance constraints. Under some assumptions, we had shown the stationary point of the approximation problem converges to a stationary point of the original problem when the parameter $\varepsilon \downarrow 0$. We proposed a sequential convex program method for solving the approximation problem and analyzed its convergence. Numerical experiments conducted on a variety of JCCDCP problems demonstrate that our method is efficient.

At each iteration, it needs to solve a large scale convex program because we use the Monte Carlo method. From our experience, neither CVX nor fmincon is a satisfactory choice to solve this type of problem rapidly and efficiently. Thus, how to solve the large scale convex program rapidly is a key factor that influences the efficiency of our method. On the other hand, many robust optimization algorithms have been proposed to avoid using the Monte Carlo method. Whether these robust optimization techniques can be applied in our method is also attractive.

Acknowledgments. X. Xiao was supported by the National Natural Science Foundation of China under project No. 11101064 and the Fundamental Research Funds for the Central Universities. L. Zhang was supported by the National Natural Science Foundation of China under project No. 11071029 and the Fundamental Research Funds for the Central Universities.

REFERENCES

[1] L. T. H. An, An efficient algorithm for globally minimizing a quadratic function under convex quadratic constraints, Math. Program., 87 (2000), 401–426.
[2] A. Charnes, W. W. Cooper and H. Symonds, Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil, Manag. Sci., 4 (1958), 235–263.
[3] Y. Gao, “Structured Low Rank Matrix Optimization Problems: A Penalty Approach,” Ph.D thesis, National University of Singapore, 2010.
[4] Y. Gao and D. Sun, Calibrating least squares semidefinite programming with equality and inequality constraints, SIAM J. Matrix Anal. Appl., 31 (2009), 1432–1457.
[5] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 1.21, April, 2011. Available from: http://cvxr.com/cvx.

[6] L. J. Hong, Y. Yang and L. Zhang, Sequential convex approximations to joint chance constrained programs: A Monte Carlo approach, Oper. Res., 59 (2011), 617–630.

[7] R. Horst and N. V. Thoni, DC programming: Overview, J. Optim. Theory Appl., 103 (1999), 1–43.

[8] D. Klatte and W. Li, Asymptotic constraint qualifications and global error bounds for convex inequalities, Math. Program., 84 (1999), 137–160.

[9] A. Nemirovski and A. Shapiro, Convex approximations of chance constrained programs, SIAM J. Optim., 17 (2006), 969–996.

[10] R. T. Rockafellar, “Convex Analysis,” Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, NJ, 1970.

[11] R. T. Rockafellar and R. J. B. Wets, “Variational Analysis,” Grundlehren der Mathemathischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 317, Springer, New York, 1998.

[12] A. Shapiro, D. Dentcheva and A. Ruszczynski, “Lectures on Stochastic Programming: Modeling and Theory,” MPS/SIAM Series on Optimization, 9, SIAM, Philadelphia, PA, 2009.

Received September 2011; 1st revision January 2012; 2nd revision March 2012.

E-mail address: xtxiao@dlut.edu.cn
E-mail address: gujian82@yahoo.cn
E-mail address: lwzhang@dlut.edu.cn
E-mail address: zhangsu@dlut.edu.cn