Partial Sample Average Approximation Method for Chance Constrained Problems

Jianqiang Cheng · **Céline Gicquel** ·
Abdel Lisser

Abstract In this paper, we present a new scheme of a sampling-based method to solve chance constrained programs. The main advantage of our approach is that the approximation problem contains only continuous variables whilst the standard sample average approximation (SAA) formulation contains binary variables. Although our approach generates new chance constraints, we show that such constraints are tractable under certain conditions. Moreover, we prove that the proposed approach has the same convergence properties as the SAA approach. Finally, numerical experiments show that the proposed approach outperforms the SAA approach on a set of tested instances.

1 Introductions

In this paper, we focus on the following chance constrained problems:

$$\min \quad f(x) \tag{1a}$$
$$(CCP) \quad s.t. \quad p_0(x) := \mathbb{P}\{g_j(x,\xi) \geq 0, j = 1, ..., m\} \geq 1 - \eta \tag{1b}$$
$$x \in X, \tag{1c}$$

where $f : R^n \to R$ is a convex objective function, $X \subset R^n$ is a convex set of additional deterministic constraints, $\xi \colon \Omega \to R^d$ is a random vector defined on a probability

Jianqiang Cheng
Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721, USA
E-mail: jqcheng@email.arizona.edu

Céline Gicquel · Abdel Lisser
Laboratoire de Recherche en Informatique (LRI),
Université Paris Sud - XI Bât. 650, 91405 Orsay Cedex, France
E-mail: {celine.gicquel, lisser}@lri.fr

space $(\Omega, \mathcal{F}, P)$ with a distribution F and its support $\Xi \subset R^d$, $g_j : R^n \times \Xi \to R$ are given constraint mapping functions, $\eta \in (0, 1]$ is a confidence parameter, and $p_0(x)$ is the feasible probability of $x \in X$. Constraint (1b) is called a chance (or probabilistic) constraint. Furthermore, constraint (1b) is called an individual chance constraint when $m = 1$ and a joint chance constraint otherwise.

Since the pioneering results of Charnes, Cooper and Symonds in [6], chance constrained optimization has attracted many researchers and practitioners due to its wide range of application areas (e.g., finance, transportation and energy [1, 8, 9, 13]). However, limited progress has been made in solving such problems for two main reasons. First, the feasible set of CCP is generally nonconvex even if the set X is convex and the function $g_j(x, \xi)$ is concave in x. Second, $\mathbb{P}\{g_j(x, \xi) \geq 0, j = 1, ..., m\}$ with a given $x \in X$ is generally hard to evaluate, as it requires multi-dimensional integrations. For instance, Nemirovski and Shapiro [18] stated that evaluating the distribution of a weighted sum of uniformly distributed independent random variables is NP-hard. For a comprehensive overview on theory and applications of chance constrained problems, we refer the reader to the books of Prékopa [21] and Shapiro et al. [25].

As chance constrained problems are computationally difficult, most computational approaches apply approximation techniques to obtain near-optimal solutions. In general, the majority of solution methods falls under two categories: convex (or tractable) approximations, and probability discretization based approaches. Nemirovski and Shapiro [18] developed a large deviation-type approximation, referred to as *Bernstein approximation*. Cheung et al. [10] established a new large deviation bounds for the stochastic linear matrix inequalities. On the other hand, probability discretization based approaches consider discrete probability distribution or representation of empirical distribution obtained from Monte Carlo samples. With regard to the discrete distribution case, Dentcheva et al. [12] introduced the concept of a p-efficient point of a probability distribution, and derived various equivalent problem formulations. Luedtke et al. [17] gave two strengthened formulations for linear programs with joint probabilistic constraints where only the right-hand side is random. Among the sampling-based approaches, there are three popular methods: scenario approach [2, 3, 5], SAA [16, 19], and constraint removal approach [4, 20].

The scenario approach (SA) was firstly introduced by Califiore and Campi [2, 3, 5], where all the sampled constraints, i.e., $g_j(x, \xi) \geq 0, j = 1, ..., m$, are satisfied for all the Monte Carlo samples of ξ. The main advantage of this approach is that a feasible solution is provided when the number of samples N is large enough. Moreover, if $g_j(x, \xi), j = 1, ..., m$ are concave in x, then the subsequent problem is a convex optimization problem. However, it is pointed out that the obtained solution is too conservative especially when η is relatively large [16]. To circumvent this drawback, an alternative sample approximation, SAA, was introduced by [16] where only a subset of sampled constraints are satisfied. When N is large enough, the SAA approach asymptotically yields accurate solutions. Moreover, the obtained solution converges to the optimal solution of the original problem under some conditions [19]. However, the resulting problem from the SAA approach is a mixed integer optimization problem with N binary variables, which is in general difficult to solve especially when N is large [19]. To avoid this drawback, the constraint removal approach [4, 20] was proposed in order to provide better solutions than SA while taking less computational

time than the SAA approach. The constraint removal approach consists in discarding a subset of sampled constraints using a heuristic algorithm, e.g., greedy algorithm, to solve the original problem.

In this paper, we present a modified sample average approximation, namely Partial Sample Average Approximation (PSAA), to solve chance constrained problems. Our approach shares the same idea as the SAA approach but we only sample a subset of random variables whereas all of the random variables are sampled in the SAA approach. The main advantage of the PSAA is its corresponding reformulation problem contains only continuous variables whilst the standard SAA formulation contains binary variables. The remainder of this paper is organized as follows. In section 2, we present a detailed study of SA and the standard SAA approach. In section 3, we present our modified SAA approach. Our general theoretical results are presented in section 4. An application of the PSAA approach to a bilinear chance constrained problem is detailed in section 5. Our numerical results are discussed in section 6. Conclusions are given in the last section.

2 Scenario approach and SAA method

When we can sample from the distribution F, the SA approximation of problem (1) can be written as

$$\min \quad f(x) \tag{2a}$$
$$s.t. \quad g_j(x, \hat{\xi}^t) \geq 0, \quad j = 1, \ldots, m, \ t = 1, \ldots, N \tag{2b}$$
$$x \in X, \tag{2c}$$

where scenarios $\hat{\xi}^1, \ldots, \hat{\xi}^N$ are independent and sampled from the distribution F. The main advantage of this approach is that there is no restriction on the distribution of ξ and it is only assumed that samples can be obtained. Moreover, when $X \in R^n$ is a convex set and g_j is concave in x for each ξ, problem (2) is a convex optimization problem. For the quality of the approximation, one main result is given by Campi et al. [5] as follows:

Theorem 1 *Assume that for all the scenario samples $\overrightarrow{\hat{\xi}} = \{\hat{\xi}^1, \ldots, \hat{\xi}^N\}$, the scenario approximation (2) is either unfeasible, or, if feasible, it attains a unique optimal solution $x^*(\overrightarrow{\hat{\xi}})$. Given $\beta \in (0, 1]$, if the sample size N satisfies the relation*

$$N \geq N^* := \lceil \frac{2}{\eta}(\log \frac{1}{\beta} + n) \rceil,$$

($\lceil \cdot \rceil$ denotes the smallest integer which is greater than or equal to the argument) then, $\mathbb{P}_{\overrightarrow{\hat{\xi}}}\{x^*(\overrightarrow{\hat{\xi}})$ is either undefined or feasible for problem (1)$\} \geq 1 - \beta$.

Theorem 1 provides a viable method to select N such that the obtained solution from SA is feasible for the original chance constrained problem. However, this approximation is too conservative as it requires that all the sampled constraints have to

be satisfied [16]. An alternative approach is *Sample Average Approximation* (SAA) where the probability distribution is replaced by its empirical distribution obtained from Monte Carlo samples of ξ. The key difference between the two approximations is that the SAA approach chooses a subset of sampled constraints to be satisfied. Accordingly, the SAA formulation of problem (1) is

$$\min \quad f(x) \tag{3a}$$

$$s.t. \quad \frac{1}{N} \sum_{t=1}^{N} \mathbb{I}(g_j(x, \hat{\xi}^t) \geq 0, \ j = 1, \ldots, m) \geq 1 - \eta \tag{3b}$$

$$x \in X, \tag{3c}$$

Where $\mathbb{I}(\mathcal{A})$ is an indicator function that takes value one when expression $\mathcal{A}$ is true and zero otherwise. When $\eta = 0$, problems (2) and (3) are equivalent. By applying the "big-M" method, the equivalent formulation of the SAA problem is as follows:

$$\min \quad f(x) \tag{4a}$$

$$s.t. \quad g_j(x, \hat{\xi}^t) + M_j^t y_t \geq 0, \ j = 1, \ldots m, \ t = 1, \ldots, N \tag{4b}$$

$$\frac{\sum_{t=1}^{N} y_t}{N} \leq \eta \tag{4c}$$

$$y_t \in \{0, 1\}, \ t = 1, \ldots, N, x \in X, \tag{4d}$$

where M_j^t is a sufficiently large constant such that constraints $g_j(x, \hat{\xi}^t) + M_j^t \geq 0, j = 1, \ldots, m, t = 1, \ldots, N$ are satisfied for all $x \in X$. It is easy to see that the new formulation is a mixed integer optimization problem. Thus, the SAA problem is in general difficult to solve when the number of samples N is large [19]. It is worth pointing out that the value of M_t may have a negative impact on the performance of the SAA reformulation. An arbitrary choice of big-M coefficients may lead to a poor LP relaxation bound, and an excess runtime to come up with an optimal solution. For more information about the impact of big-M coefficients on the performance of SAA, we refer the reader to [26].

3 Partial Sample Average Approximation

Before presenting our approach, we first introduce a key assumption which the prososed approach will rely on.

Assumption 1. *We assume that $\xi = (\xi_1, \xi_2)$ where ξ_1 and ξ_2 are independent.*

Without loss of generality, the dimensions of vectors ξ_1 and ξ_2 are assumed to be d_1 and d_2 respectively. It is easy to see that the dimension of ξ is $d = d_1 + d_2$.

Lemma 1 *Let X and Y be independent integrable random variables and $g(x, y)$ be a real-valued function. If the expectation of $g(X, Y)$ exists, then*

$$\mathbb{E}[g(X, Y)] = \mathbb{E}_X \mathbb{E}_Y[g(X, Y)] = \mathbb{E}_Y \mathbb{E}_X[g(X, Y)]. \tag{5}$$

We reconsider the chance constraint (1b) with Assumption 1. First, the probability is represented as an expectation:

$$p_0(x) = \mathbb{P}\{G(x, \xi_1, \xi_2) \geq 0\} = \mathbb{E}[\mathbb{I}(G(x, \xi_1, \xi_2) \geq 0)],$$

where $G(x, \xi_1, \xi_2) := \min_{1 \leq j \leq m} g_j(x, \xi_1, \xi_2)$. By Lemma 1, the chance constraint (1b) is equivalent to the following constraint:

$$p_0(x) = \mathbb{E}_{\xi_1}(\mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \xi_1, \xi_2) \geq 0)]) \geq 1 - \eta.$$

We apply the same idea as the SAA method, replacing the distribution of ξ_1 by its empirical distribution based on N Monte Carlo samples. More precisely,

$$\begin{aligned}
\mathbb{E}_{\xi_1}(\mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \xi_1, \xi_2) \geq 0)]) &\approx \frac{\sum_{t=1}^{N} \mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) \geq 0)]}{N} \\
&= \frac{\sum_{t=1}^{N} \mathbb{P}\{G(x, \hat{\xi}_1^t, \xi_2) \geq 0\}}{N}
\end{aligned}$$

where $\hat{\xi}_1^1, ..., \hat{\xi}_1^N$ are independent Monte Carlo samples of random vector ξ_1.

Definition 1 (**Partial sample average approximation**) Let $\hat{\xi}_1^1, ..., \hat{\xi}_1^N$ be independent Monte Carlo samples of random vector ξ_1. Then, we have a partial sample average approximation of problem (CCP) as follows:

$$\min \quad f(x) \tag{8a}$$
$$s.t. \quad p_N(x) := \frac{\sum_{t=1}^{N} \mathbb{P}\{G(x, \hat{\xi}_1^t, \xi_2) \geq 0\}}{N} \geq 1 - \eta \tag{8b}$$
$$x \in X, \tag{8c}$$

which is equivalent to

$$f_{PSAA}^* := \min \quad f(x) \tag{9a}$$
$$s.t. \quad \mathbb{P}\{g_j(x, \hat{\xi}_1^t, \xi_2) \geq 0, j = 1, ..., m\} \geq y_t, \ t = 1, ..., N \tag{9b}$$
$$(PSAA) \quad \frac{\sum_{t=1}^{N} y_t}{N} \geq 1 - \eta \tag{9c}$$
$$y_t \geq 0; t = 1, ..., N, x \in X, \tag{9d}$$

When compared to the SAA approach, the main advantage of the PSAA method is that the resulting equivalent problem, i.e., problem (9), contains only continuous variables but at the cost of introducing a set of new chance constraints. However, we can show that the newly generated constraints could be easily handled under certain conditions as illustrated in section 5.

4 Our theoretical results

4.1 Convexity results of PSAA

It is well known that the feasible set of problem (1) is generally nonconvex even if the set X is convex and the functions $g_j(x, \xi), j = 1 \ldots, m$ are concave in x. However, there exists some convexity results in the literature for chance constrained problems under some given assumptions. For instance, individual chance constraints with normally distributed random variables can be reformulated as second order cone constraints when $\eta \leq 0.5$.

Prior to presenting our convexity results, we first define the concave points for probability distributions and then show examples for some univariate distributions.

Definition 2 (**Concave point**) z^* is said to be a *concave point* of a random variable ξ, if the cumulative distribution function $\Phi(z)$ of ξ is concave for all $z \geq z^*$.

Evidently, if z^* is a concave point of ξ, then z' is also a concave point provided $z' \geq z^*$. When ξ is univariate and has the standard normal distribution, then 0 is a concave point of ξ and it is also the minimal concave point. Table 1 lists the minimal concave points for some other well-known univariate distributions.

Table 1 Table of selected univariate distributions

Distribution	Density function	z^*
Normal	$\frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(z-\mu)^2}{2\sigma^2})$	μ
Exponential	$\lambda \exp(-\lambda z) \quad z \geq 0$	0
Uniform	$\frac{1}{b-a} \quad a \leq z \leq b$	a
Weibull	$abz^{b-1} \exp(-az^b) \quad z \geq 0$	$0 \qquad 0 < b \leq 1$ $(\frac{b-1}{ab})^{\frac{1}{b}} \quad 1 < b$
Gamma	$\frac{z^{k-1}e^{-\frac{z}{\theta}}}{\theta^k \Gamma(k)} \quad z > 0$	$0 \qquad 0 < k \leq 1$ $(k-1)\theta \quad 1 < k$
Student	$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\,\Gamma(\frac{\nu}{2})}\left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	0

Although Table 1 shows the minimal concave points for some popular distributions, it is still difficult to find z^* for general multivariate distributions. To come up with the concave point of multivariate normal distribution, we use the following relevant result:

Theorem 2 *[22, 23] $\Phi(z_1, ..., z_d; \Sigma)$ is concave in the set $\{z | z_i \geq \sqrt{d-1}, i = 1, ..., d\}$, where $\Phi(z_1, ..., z_d; \Sigma)$ is the d-variate standard normal probability distribution function with correlation matrix Σ.*

Note that $(\sqrt{d-1}, ..., \sqrt{d-1})$ is a concave point of d-variate standard normally distributed variables for any correlation matrix Σ. But, it is still challenging to find the minimal (or Pareto) concave point for the normal distribution.

We reconsider the chance constraint (8b) with the following specific form:

$$\mathbb{P}\{g_j(x, \xi_1^t) \geq \xi_{2_j}, \quad j = 1, \ldots, m\} \geq 1 - \eta. \tag{10}$$

Theorem 3 *Assume that functions $g_j(x, \hat{\xi}_1^t)$ $j = 1, \ldots, m$ are concave in x and there exists a concave point $z^* = (z_1^*, ..., z_m^*)$ of $\xi_2 = (\xi_{2_1}, ..., \xi_{2_m})$ such that $g_j(x, \hat{\xi}_1^t) \geq z_j^*$ holds for any $x \in X$. Then the following set is convex*

$$X_0(\xi_1^t) := \{(x; y_t) \in R^{n+1} : \mathbb{P}\{g_j(x, \hat{\xi}_1^t) \geq \xi_{2_j}, \quad j = 1, \ldots, m\} \geq y_t, x \in X\}. \tag{11}$$

Therefore the feasible set of the corresponding PSAA problem is convex.

Proof. Let $G_0(x) := \mathbb{P}\{g_j(x, \hat{\xi}_1^t) \geq \xi_{2_j}, j = 1, \ldots, m\} = \Phi(g_1(x, \hat{\xi}_1^t), ..., g_m(x, \hat{\xi}_1^t))$ where $\Phi(z)$ is the cumulative distribution function of ξ_2.

We first prove that $G_0(x)$ is concave on X. For any two points $x_1, x_2 \in X$ and $\lambda \in [0, 1]$, we have

$$
\begin{aligned}
G_0(\lambda x_1 + (1-\lambda)x_2) &= \Phi\Big(g_1(\lambda x_1 + (1-\lambda)x_2, \hat{\xi}_1^t), ..., g_m(\lambda x_1 + (1-\lambda)x_2, \hat{\xi}_1^t)\Big) \\
&\geq \Phi\Big(\lambda g_1(x_1, \hat{\xi}_1^t) + (1-\lambda)g_1(x_2, \hat{\xi}_1^t), ..., \lambda g_m(x_1, \hat{\xi}_1^t) + (1-\lambda)g_m(x_2, \hat{\xi}_1^t)\Big) \\
&\geq \lambda \Phi\Big(g_1(x_1, \hat{\xi}_1^t), ..., g_m(x_1, \hat{\xi}_1^t)\Big) + (1-\lambda)\Phi\Big(g_1(x_2, \hat{\xi}_1^t), ..., g_m(x_2, \hat{\xi}_1^t)\Big) \\
&= \lambda G_0(x_1) + (1-\lambda)G_0(x_2)
\end{aligned}
$$

where the first inequality is due to the concavity of $g_j(x, \hat{\xi}_1^t)$ and the fact that $\Phi(z)$ is a non-decreasing function. The second inequality is due to the concavity of $\Phi(z)$ when $z \geq z^*$ as well as the assumption that $g_j(x, \hat{\xi}_1^t) \geq z_j^*$ holds for any $x \in X$. Thus $G_0(x)$ is concave on X and consequently $X_0(\xi_1^t)$ is convex, which concludes the proof. $\qquad\square$

Example

We consider a simple example to illustrate the results of Theorem 3:

$$\min \quad c^T x \tag{12a}$$
$$s.t. \quad \mathbb{P}\{\xi_1^T x \geq \xi_2\} \geq p \tag{12b}$$
$$x \geq 0 \tag{12c}$$

where $c \in R^n$ is a deterministic cost vector, $\xi_1 \in R^n$ and $\xi_2 \in R$ are random, and both are assumed uniformly distributed. If the lower bound of ξ_2 is 0 and the lower bound of ξ_1 is nonnegative, then the corresponding PSAA problem is a convex problem, as the minimal concave point of ξ_2 is 0. It is worth noting that problem (12) is NP-hard [18] when ξ_1 is independent uniformly distributed in a box.

4.2 PSAA vs SAA

In general, PSAA and SAA have the same convergence properties as they share the same idea: replacing the probability distribution by its empirical distribution obtained from Monte Carlo samples.

In the following, we present our PSAA convergence results. First of all, we denote by f^*, f^*_{SAA} and f^*_{PSAA} the optimal values of the original problem, SAA and PSAA problems respectively. S, S_{SAA} and S_{PSAA} are the sets of the optimal solutions defined accordingly. We first recall two definitions.

Definition 3 Functions f_N are said to epiconverge to f w.p.1, denoted by $f_N \xrightarrow{e} f$ w.p.1, if for almost every (a.e.) $\omega \in \Omega$ the functions $f_N(\cdot, \omega)$ epiconverge to $f(\cdot)$.

Definition 4 Function $g(x, \xi)$ is a *Carathéodory function*, i.e., $g(x, \xi)$ is measurable for every $x \in R^n$ and $g(x, \xi)$ is continuous for a.e. $\xi \in \Xi$.

Lemma 2 *Let $G(x, \xi_1, \xi_2) := \min_{1 \leq j \leq m} g_j(x, \xi_1, \xi_2)$ be a* Carathéodory *function. Then $p_N(x) \xrightarrow{e} p_0(x)$ w.p.1. Furthermore, suppose that the set X is compact, the objective function $f(x)$ is continuous, and the following condition holds: (a) there is an optimal solution x^* of problem (1), such that for any $\epsilon > 0$, there is $x \in X$ with $\|x - x^*\| \leq \epsilon$ and $p_0(x) > 1 - \eta$, . Then $f^*_{PSAA} \to f^*$ and $\mathbf{D}(S_{PSAA}, S) \to 0$ w.p.1 as $N \to \infty$, where $\mathbf{D}(A, B) := \sup_{x \in A} dist(x, B)$ is the deviation of set A from set B.*

Proof. Let $\bar{p}_0(x) = 1 - p_0(x)$, $\bar{p}_N(x) = 1 - p_N(x)$, $\bar{S}$ and $\bar{S}_{PSAA}$ be the complements of the sets S and S_{PSAA}, respectively. We first prove that $\bar{p}_N(x)$ is lower semicontinuous. Since the indicator function on the open set is semicontinuous and $G(x, \hat{\xi}_1^t, \xi_2)$ is a *Carathéodory* function, then $\mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) < 0)$ is random lower semicontinuous by Proposition 12.45 in [24]. Thus, by Fatou's lemma we have for any x_0,

$$\liminf_{x \to x_0} \mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) < 0)] = \liminf_{x \to x_0} \int_{\Xi_2} \mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) dP(\xi_2) \tag{13}$$

$$\geq \int_{\Xi_2} \liminf_{x \to x_0} \mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) dP(\xi_2) \tag{14}$$

$$\geq \int_{\Xi_2} \mathbb{I}(G(x_0, \hat{\xi}_1^t, \xi_2) dP(\xi_2) = \mathbb{E}_{\xi_2}[\mathbb{I}(G(x_0, \hat{\xi}_1^t, \xi_2) < 0)], \tag{15}$$

where the second inequality holds due to the lower semicontinuity of $\mathbb{I}(G(x_0, \hat{\xi}_1^t, \xi_2)$. This shows $\mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) < 0)]$ is lower semicontinuous. Analogously, we can show $\bar{p}_N(x) = 1 - p_N(x) = \frac{\sum_{t=1}^{N} \mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) < 0)]}{N}$ is also lower semicontinuous. As $0 \leq \mathbb{E}_{\xi_2}[\mathbb{I}(G(x, \hat{\xi}_1^t, \xi_2) < 0)] \leq 1$ and N samples are independent and identically distributed, then we have $\bar{p}_N(x) \xrightarrow{e} \bar{p}_0(x)$ w.p.1 based on the results of Theorem 7.51 in [25]. Therefore, $p_N(x) \xrightarrow{e} p_0(x)$ w.p.1.

The remaining of the proof follows the same procedure of Proposition 5.30 in [25] (or the proof of Proposition 2.2 in [19]). By condition (a), the set S is nonempty and there is $\bar{x}$ such that $p_0(\bar{x}) > 1 - \eta$ and $\bar{p}_0(\bar{x}) < \eta$. Moreover, we have $\bar{p}_N(\bar{x})$ converges to $\bar{p}_0(\bar{x})$ w.p.1. Subsequently, $\bar{p}_N(\bar{x}) < \eta$, and thus the PSAA problem (i.e.,

problem (8)) has a feasible solution, w.p.1 for N large enough. Since $\bar{p}_N(x)$ is lower semicontinuous, the feasible set of the PSAA problem is compact, and thus S_{PSAA} is nonempty w.p.1 for N large enough. As $\bar{x}$ is feasible for the PSAA problem, then $f(\bar{x}) \geq f^*_{PSAA}$. Because of the continuity of $f(x)$ and condition (a), for any $\epsilon > 0$, there exists $\bar{x} \in X$ sufficiently close to x^* such that $\bar{p}_0(\bar{x}) < \eta$ and $f(\bar{x}) \leq f(x^*) + \epsilon = f^* + \epsilon$. Thus, we have

$$\limsup_{N \to \infty} f^*_{PSAA} \leq f(x^*) + \epsilon = f^* + \epsilon, w.p.1.$$

Since $\epsilon > 0$ is arbitrary, then we have

$$\limsup_{N \to \infty} f^*_{PSAA} \leq f^*, w.p.1.$$

Let $\bar{x}_N \in S_{PSAA}$, i.e., $\bar{x}_N \in X$, $\bar{p}_N(\bar{x}_N) \leq \eta$ and $f^*_{PSAA} = f(\bar{x}_N)$. As the set X is compact, we can assume by passing to a subsequence if necessary that $\bar{x}_N$ converges to a point $\hat{x} \in X$ w.p.1. Also we have $\bar{p}_N(x) \xrightarrow{e} \bar{p}_0(x)$ w.p.1, and hence

$$\liminf_{N \to \infty} \bar{p}_N(\bar{x}_N) \geq \bar{p}_0(\hat{x}), w.p.1.$$

It follows that $\bar{p}_0(\hat{x}) \leq \eta$ and thus $\hat{x}$ is a feasible solution of the original problem. Subsequently, $f(\hat{x}) \geq f^*$. Also $f(\bar{x}_N) \to f(\hat{x}), w.p.1$, and hence

$$\liminf_{N \to \infty} f^*_{PSAA} \geq f^*, w.p.1.$$

Thus, we conclude that $f^*_{PSAA} \to f^*, w.p.1$. Moreover, it follows that the point $\hat{x}$ is an optimal solution of the original problem and consequently we have $\mathbf{D}(S_{PSAA}, S) \to 0$ $w.p.1$. □

5 Bilinear chance constraints

In this section, we present a special case of problem (1) where its subsequent PSAA approximation is tractable. We consider the following bilinear chance constrained problem:

$$\min \quad f(x) \tag{16a}$$
$$s.t. \quad \mathbb{P}\{M(x)\xi_1 \geq \xi_2 \mathbf{e}_m\} \geq 1 - \eta \tag{16b}$$
$$x \in X \tag{16c}$$

where $M(x) \in R^{m \times d_1}$ is an affine matrix of x, $\mathbf{e}_m \in R^m$ is all-ones vector, $\xi_1 \in R^{d_1}$ and $\xi_2 \in R$ are random variables. The bilinear chance constrained problem is widely applied in a broad range of areas, such as finance [15], production planning problems [7], knapsack problems [26], and energy [14].

Lemma 3 *The PSAA formulation of problem* (16) *is equivalent to the following problem:*

$$\min \quad f(x) \tag{17a}$$

$$s.t. \quad M(x)\hat{\xi}_1^t \geq \Phi^{-1}(y_t)\mathbf{e}_m, \ t = 1, ..., N \tag{17b}$$

$$\frac{\sum_{t=1}^N y_t}{N} \geq 1 - \eta \tag{17c}$$

$$0 \leq y_t \leq 1; t = 1, ..., N \tag{17d}$$

$$x \in X \tag{17e}$$

where $\Phi(z)$ is the cumulative distribution function of ξ_2 and $\Phi^{-1}(p)$ is the generalized inverse distribution function defined as: $\Phi^{-1}(p) = \inf\{y \in R : \Phi(y) \geq p\}$.

Proof. Based on problem (9), the approximation of the chance constraint (16b) becomes:

$$\mathbb{P}\{M(x)\hat{\xi}_1^t \geq \xi_2\mathbf{e}_m\} \geq y_t, \frac{\sum_{t=1}^N y_t}{N} \geq 1 - \eta \ y_t \geq 0; t = 1, ..., N.$$

By the definition of $\Phi^{-1}(p)$, $\mathbb{P}\{M(x)\hat{\xi}_1^t \geq \xi_2\mathbf{e}_m\} \geq y_t$ is equivalent to $M(x)\hat{\xi}_1^t \geq \Phi^{-1}(y_t)\mathbf{e}_m$, which completes the proof. □

Lemma 3 shows that with the PSAA method, the chance constraint is approximated by constraints (17b)-(17d) where only constraints (17b) are not linear due to the inverse distribution function $\Phi^{-1}(y_t)$. If $\Phi^{-1}(y_t)$ can be approximated by a convex function, then we have a convex approximation for the PSAA problem. As a matter of illustration, we present a linear approximation for $\Phi^{-1}(y_t)$ when ξ_2 is uniformly distributed on the interval $[L, U]$, specifically $\Phi^{-1}(y_t) \approx y_t(U - L) + L$. Note that $\Phi^{-1}(0) = -\infty$.

Theorem 4 *If ξ_2 is uniformly distributed on the interval $[L, U]$, then the corresponding PSAA problem of the bilinear problem has a conservative (or safe) approximation as follows:*

$$\min \quad f(x) \tag{18a}$$

$$s.t. \quad \frac{M(x)\hat{\xi}_1^t - L\mathbf{e}_m}{U - L} \geq y_t\mathbf{e}_m, \ t = 1, ..., N \tag{18b}$$

$$\frac{\sum_{t=1}^N y_t}{N} \geq 1 - \eta \tag{18c}$$

$$y_t \leq 1; t = 1, ..., N \tag{18d}$$

$$x \in X. \tag{18e}$$

Moreover, if there is an optimal solution of PSAA problem denoted by x^{PSAA} satisfying constraints $M(x^{PSAA})\hat{\xi}_1^t \geq L\mathbf{e}_m$ for all samples $\hat{\xi}_1^t, t = 1, ..., N$, then x^{PSAA} is also an optimal solution of the conservative problem and thus the conservative approximation becomes equivalent to the PSAA problem.

Proof. Let (x^{fea}, y^{fea}) be a feasible solution of problem (18). With the PSAA method, the chance constraint (16b) is approximated by the following constraint:

$$\frac{\sum_{t=1}^{N} \mathbb{P}\{M(x)\hat{\xi}_1^t \geq \xi_2 \mathbf{e}_m\}}{N} \geq 1 - \eta \tag{19}$$

Let $z_t = \mathbb{P}\{M(x^{fea})\hat{\xi}_1^t \geq \xi_2 \mathbf{e}_m\}$. Since ξ_2 is uniformly distributed on the interval $[L, U]$, we have:

$$z_t = \begin{cases} 0 & \text{if } \min\{M(x^{fea})\hat{\xi}_1^t\} < L \\ \frac{\min\{M(x^{fea})\hat{\xi}_1^t\} - L}{U - L} & \text{if } L \leq \min\{M(x^{fea})\hat{\xi}_1^t\} \leq U \\ 1 & \text{if } U \leq \min\{M(x^{fea})\hat{\xi}_1^t\} \end{cases} \tag{20}$$

As (x^{fea}, y^{fea}) is a feasible solution of problem (18), we have $y_t^{fea} \leq 1$, $\frac{\sum_{t=1}^{N} y_t^{fea}}{N} \geq 1 - \eta$ and $\frac{M(x^{fea})\hat{\xi}_1^t - L\mathbf{e}_m}{U - L} \geq y_t^{fea}\mathbf{e}_m$. Thus, we conclude that $z_t \geq y_t$, which leads to the conclusion that $\frac{\sum_{t=1}^{N} z_t}{N} \geq 1 - \eta$. Therefore, x^{fea} is also a feasible solution of the PSAA problem, and problem (18) is a conservative approximation of the PSAA problem.

Let x^{PSAA} be an optimal solution of the PSAA problem which satisfies the constraint $M(x^{PSAA})\hat{\xi}_1^t \geq L\mathbf{e}_m$. Let $\pi_t = \mathbb{P}\{M(x^{PSAA})\hat{\xi}_1^t \geq \xi_2 \mathbf{e}_m\}$. Then we have $\pi_t \leq \min\{1, \frac{\min\{M(x^{PSAA})\hat{\xi}_1^t\} - L}{U - L}\}$ and $\frac{\sum_{t=1}^{N} \pi_t}{N} \geq 1 - \eta$. Thus $(x^{PSAA}, \pi_1, ..., \pi_N)$ is a feasible solution of problem (18). Moreover, problem (18) is a conservative approximation of the PSAA problem. Therefore, $(x^{PSAA}, \pi_1, ..., \pi_N)$ is an optimal solution of problem (18). $\square$

6 Numerical experiments

In this section, we consider two different bilinear chance constrained problems to evaluate the performance of our PSAA approach: a simple example and a supply/demand equilibrium problem. All the problems are solved using CPLEX 12.8 [11] with its default parameters on a machine with an Intel(R) Core(TM) i7-4600U @ 2.10 GHz 2.70 GHz with 16.0 GB RAM.

6.1 A simple example

Consider the aforementioned simple example:

$$\min \quad c^T x \tag{21a}$$
$$s.t. \quad \mathbb{P}\{\xi_1^T x \geq \xi_2\} \geq 1 - \eta \tag{21b}$$
$$x \geq 0, \tag{21c}$$

where $c \in R^n$ is a deterministic cost vector, $\xi_1 \in R^n$ and $\xi_2 \in R$ are independent uniformly distributed random vector and variable repectively. As we mentioned in section 4, problem (21) is NP-hard [18] when ξ_1 is independent uniformly distributed in a box.

The parameters are set as follows: $n = 50$, $\eta \in \{0.05, 0.1, 0.15, 0.2\}$, and c is uniformly generated on the interval $[10, 100]$. The upper and lower bounds of ξ_1 are uniformly generated on the intervals $[10, 20]$ and $[5, 10]$, respectively, whilst the upper and lower bounds of ξ_2 are uniformly generated on the intervals $[50, 100]$ and $[0, 50]$, respectively.

We compare our method with the SAA approach with $M_t = \hat{\xi}_2^t$ and SA. The results are reported in Table 2 where, for each value of η, row 1 gives the number of scenarios, rows 2, 3 and 4 present the objective value, the CPU time and the out-of-sample probability [1] for SA, SAA and PSAA, respectively.

	$\eta = 0.05$			$\eta = 0.10$		
	SA	SAA	PSAA	SA	SAA	PSAA
N=500						
Obj	123.85	99.79	100.42	123.85	90.65	92.60
CPU(S)	0.04	0.67	0.10	0.04	1.70	0.13
Prob	0.994	0.941	0.946	0.994	0.883	0.898
N=1000						
Obj	124.65	100.22	100.98	124.65	92.07	92.75
CPU(S)	0.06	7.39	0.30	0.06	34.88	0.41
Pro	0.996	0.942	0.948	0.996	0.892	0.900
N=10000						
Obj	131.84	–	101.29	131.84	–	92.77
CPU(S)	0.38	–	1.80	0.38	–	1.70
Prob	0.999	–	0.950	0.999	–	0.902
	$\eta = 0.05$			$\eta = 0.10$		
	SA	SAA	PSAA	SA	SAA	PSAA
N=500						
Obj	123.85	85.57	86.57	123.85	79.63	81.34
CPU(S)	0.04	2.69	0.22	0.04	4.03	0.15
Prob	0.994	0.836	0.847	0.994	0.772	0.797
N=1000						
Obj	124.65	85.80	86.66	124.65	80.64	81.48
CPU(S)	0.06	80.04	0.53	0.06	70.81	0.50
Prob	0.996	0.841	0.849	0.996	0.788	0.799
N=10000						
Obj	131.84	–	86.83	131.84	–	81.72
CPU(S)	0.38	–	1.65	0.38	–	1.65
Prob	0.999	–	0.851	0.999	–	0.8013

Table 2 Computational results of the simple example when $n = 50$. "–" indicates no optimal solutions found within two hours.

From Table 2, we observe that SA requires less runtime than both SAA and PSAA approaches. However, the solutions provided by SA are significantly more expensive

[1] It is the empirical feasible probability of given optimal solution x^* attained from different approximation approaches with 100000 scenarios.

than the solutions provided by the SAA and PSAA approaches. For instance, when $\eta = 0.10$ and $N = 1000$, the solution of SA is 34% more expensive than the ones of both SAA and PSAA approaches. Notice that between SAA and PSAA approaches, the PSAA solution cost is slightly higher than the SAA solution cost. But compared with the SAA approach, the PSAA solutions provide improved empirical probabilities, as they are closer to the target threshold of $1 - \eta$. For instance, when $\eta = 0.10$ and $N = 1000$, the cost of the PSAA solution is 0.74% higher than the cost of the SAA approach but its empirical probability is the same as the target level. Moreover, the PSAA approach takes less runtime than the SAA approach especially for $N = 10000$ where the SAA approach solved no instance within two hours for all the four values of η whilst the PSAA approach solved any instance within 2 seconds.

6.2 Problem of supply/demand equilibrium under uncertainty

In this section, we consider a supply-demand equilibrium problem [14]. This problem is commonly used in energy management problems and is a sub-problem of the Unit Commitment Problem (UCP). It aims at minimizing the global production cost while ensuring the supply-demand balance and satisfying the operational constraints of a mix of power generation units (hydraulic valleys, nuclear plants and classical thermal units -coal, fuel and gas-) for a discrete time-period planning horizon.

We propose the following concise formulation that emphasizes the structure of the problem :

$$
\begin{aligned}
&\min \; \sum_{i=1}^{n} \sum_{t=1}^{T} c_{i,t} x_{i,t} \\
&\text{s.t.} \; \mathbb{P}\{\sum_{i=1}^{n} \tilde{A}_{i,t} x_{i,t} \geq \tilde{D}_t = \sum_{j=1}^{m} b_{t,j} \tilde{D}_j, \; t = 1, ..., T\} \geq 1 - \eta \\
&\quad\quad \sum_{t=1}^{T} x_{i,t} \leq r_i T, \; i = 1, ..., n \\
&\quad\quad 0 \leq x_{i,t} \leq 1, \; i = 1, ..., n, \; t = 1, ..., T
\end{aligned}
\tag{22}
$$

where

- $c_{i,t}$ is the production cost for plant i at time step t;
- r_i is the maximum proportion that plant i will be used over the planning horizon. These constraints may represent planned shutdown for maintenance operations for instance;
- $x_{i,t}$ is the control variable for plant i at time step t ;
- $x_t = (x_{1,t}, ..., x_{n,t}), \; t = 1, ..., T$;
- $\tilde{A}_t = (\tilde{A}_{1,t}, ..., \tilde{A}_{n,t})$ is a random vector representing the availability of the production units at time step t;
- $\tilde{D}_t$ is a random variable representing the total demand at time step t;
- $(\tilde{D}_1, ..., \tilde{D}_m)$ is a random vector on which $\tilde{D}_t$ depends linearly whereas $b_{t,j}$ is the coefficient.

For the sake of simplicity, we assume that $\tilde{A}_t$ and $\tilde{D}_j$ are independent uniformly distributed. The lower and upper bounds of $\tilde{A}_{i,t}$ are uniformly generated on the intervals $[20, 60]$ and $[60, 110]$, respectively. The lower and upper bounds of $\tilde{D}_j$ are uniformly generated on the intervals $[10, 30]$ and $[30, 50]$, respectively. The other parameters are set as follows: $n = 50$, $T = 10$, $m = 10$, c is uniformly generated

on the interval $[0, 100]$, r is uniformly generated on the interval $[0.6, 1]$, and $b_{t,j}$ is uniformly generated on the interval $[1, 3]$. Moreover, four confidence parameters are considered: $\eta \in \{0.05, 0.1, 0.15, 0.2\}$. We also compare our method with the SA and the SAA approaches with $M_\tau^t = \hat{D}_\tau^t$, where $\hat{D}_\tau^t$ is the t-th scenario of $\tilde{D}_\tau$.

Numerical results are given by Table 3 where the columns show the same information results as Table 2. From Table 3, we can draw similar conclusions regarding the efficacy of the PSAA approach. The latter finds less conservative solutions than SA, and is significantly faster than the SAA approach. For instance, the SAA approach failed to solve any instance within two hours whilst the PSAA approach solved the largest instance, i.e., $N = 10000$, within 59 seconds.

	$\eta = 0.05$			$\eta = 0.10$		
	SA	SAA	PSAA	SA	SAA	PSAA
N=500						
Obj	2432.3	–	2338.9	2432.3	–	2246.6
CPU(S)	0.06	–	0.24	0.06	–	0.25
Prob	0.968	–	0.930	0.968	–	0.878
N=1000						
Obj	2538.8	–	2339.2	2538.8	–	2238.3
CPU(S)	0.13	–	0.58	0.13	–	0.59
Pro	0.975	–	0.937	0.975	–	0.896
N=10000						
Obj	2768.6	–	2356.5	2768.6	–	2252.3
CPU(S)	1.85	–	31.02	1.85	–	53.15
Prob	0.995	–	0.949	0.995	–	0.909
	$\eta = 0.15$			$\eta = 0.20$		
	SA	SAA	PSAA	SA	SAA	PSAA
N=500						
Obj	2432.3	–	2181.8	2432.3	–	2131.4
CPU(S)	0.06	–	0.25	0.06	–	0.26
Prob	0.968	–	0.836	0.968	–	0.794
N=1000						
Obj	2538.8	–	2172.0	2538.8	–	2121.2
CPU(S)	0.13	–	0.61	0.13	–	0.67
Prob	0.975	–	0.845	0.975	–	0.801
N=10000						
Obj	2768.6	–	2184.1	2768.6	–	2131.5
CPU(S)	1.85	–	50.32	1.85	–	58.94
Prob	0.995	–	0.868	0.995	–	0.824

Table 3 Computational results of supply/demand problem $n = 50$. "–" indicates indicates no optimal solutions found within two hours.

7 Conclusions

In this paper, we present a new variant of sample average approximation (SAA) method for chance constrained problems called partial sample average approximation (PSAA). We show that the PSAA approach has similar convergence properties as the standard SAA approach, and the corresponding reformulation problem is convex under certain conditions while the standard SAA approach yields a mixed integer problem. In addition, we introduce a new tool we call concave points which plays an important role in the convexity of our approximation. Although the PSAA approach relies on the independence assumption and generates new chance constraints, our numerical results show its significant improvements both in terms of the solution quality and runtimes when compared to SA and the SAA approach, respectively.

References

1. Andrieu, L., Henrion, R., Römisch, W.: A model for dynamic chance constraints in hydro power reservoir management. European Journal of Operational Research **207**(2), 579–589 (2010)
2. Calafiore, G., Campi, M.C.: Uncertain convex programs: randomized solutions and confidence levels. Mathematical Programming **102**(1), 25–46 (2005)
3. Calafiore, G.C., Campi, M.C.: The scenario approach to robust control design. Automatic Control, IEEE Transactions on **51**(5), 742–753 (2006)
4. Campi, M.C., Garatti, S.: A sampling-and-discarding approach to chance-constrained optimization: feasibility and optimality. Journal of Optimization Theory and Applications **148**(2), 257–280 (2011)
5. Campi, M.C., Garatti, S., Prandini, M.: The scenario approach for systems and control design. Annual Reviews in Control **33**(2), 149–157 (2009)
6. Charnes, A., Cooper, W.W., Symonds, G.H.: Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. Management Science **4**(3), 235–263 (1958)
7. Cheng, J., Gicquel, C., Lisser, A.: A second-order cone programming approximation to joint chance-constrained linear programs. In: International Symposium on Combinatorial Optimization, pp. 71–80. Springer (2012)
8. Cheng, J., Lisser, A.: A second-order cone programming approach for linear programs with joint probabilistic constraints. Operations Research Letters **40**(5), 325–328 (2012)
9. Cheng, J., Lisser, A.: Maximum probability shortest path problem. Discrete Applied Mathematics **192**, 40–48 (2015)
10. Cheung, S.S., Man-Cho So, A., Wang, K.: Linear matrix inequalities with stochastically dependent perturbations and applications to chance-constrained semidefinite optimization. SIAM Journal on Optimization **22**(4), 1394–1430 (2012)
11. CPLEX, I.I.I.: High-performance mathematical programming engine. International Business Machines Corp (2010)
12. Dentcheva, D., Prékopa, A., Ruszczynski, A.: Concavity and efficient points of discrete distributions in probabilistic programming. Mathematical Programming **89**(1), 55–77 (2000)
13. Dentcheva, D., Ruszczyński, A.: Portfolio optimization with stochastic dominance constraints. Journal of Banking & Finance **30**(2), 433–451 (2006)

14. Gorge, A.: Programmation semi-définie positive. méthodes et algorithmes pour le management d'énergie. Ph.D. thesis, Université Paris Sud-Paris XI (2013)
15. Kogan, A., Lejeune, M.A.: Threshold boolean form for joint probabilistic constraints with random technology matrix. Mathematical Programming **147**(1-2), 391–427 (2014)
16. Luedtke, J., Ahmed, S.: A sample approximation approach for optimization with probabilistic constraints. SIAM Journal on Optimization **19**(2), 674–699 (2008)
17. Luedtke, J., Ahmed, S., Nemhauser, G.L.: An integer programming approach for linear programs with probabilistic constraints. Mathematical Programming **122**(2), 247–272 (2010)
18. Nemirovski, A., Shapiro, A.: Convex approximations of chance constrained programs. SIAM Journal on Optimization **17**(4), 969–996 (2006)
19. Pagnoncelli, B., Ahmed, S., Shapiro, A.: Sample average approximation method for chance constrained programming: theory and applications. Journal of optimization theory and applications **142**(2), 399–416 (2009)
20. Pagnoncelli, B.K., Reich, D., Campi, M.C.: Risk-return trade-off with the scenario approach in practice: a case study in portfolio selection. Journal of Optimization Theory and Applications **155**(2), 707–722 (2012)
21. Prékopa, A.: Stochastic programming. Springer (1995)
22. Prékopa, A.: On the concavity of multivariate probability distribution functions. Operations research letters **29**(1), 1–4 (2001)
23. Prékopa, A.: Probabilistic programming. Handbooks in operations research and management science **10**, 267–351 (2003)
24. Rockafellar, R.T., Wets, R.J.B.: Variational analysis, vol. 317. Springer Science & Business Media (2009)
25. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on stochastic programming: modeling and theory, vol. 9. SIAM (2009)
26. Song, Y., Luedtke, J.R., Küçükyavuz, S.: Chance-constrained binary packing problems. INFORMS Journal on Computing **26**(4), 735–747 (2014)