

Validation Analysis of Mirror Descent Stochastic Approximation Method

Guanghui Lan · Arkadi Nemirovski · Alexander Shapiro

Received: date / Accepted: date

Abstract The main goal of this paper is to develop accuracy estimates for stochastic programming problems by employing stochastic approximation (SA) type algorithms. To this end we show that while running a Mirror Descent Stochastic Approximation procedure one can compute, with a small additional effort, lower and upper statistical bounds for the optimal objective value. We demonstrate that for a certain class of convex stochastic programs these bounds are comparable in quality with similar bounds computed by the sample average approximation method, while their computational cost is considerably smaller.

Keywords stochastic approximation, sample average approximation method, stochastic programming, Monte Carlo sampling, mirror descent algorithm, prox-mapping, optimality bounds, large deviations estimates, asset allocation problem, conditional value-at-risk.

1 Introduction

Consider the following Stochastic Programming (SP) problem

$$\text{Opt} = \min_{x \in X} \{f(x) := \mathbb{E}[F(x, \xi)]\}, \quad (1.1)$$

where $X \subset \mathbb{R}^n$ is a nonempty bounded closed convex set, ξ is a random vector whose probability distribution P is supported on set $\Xi \subset \mathbb{R}^d$ and $F : X \times \Xi \rightarrow \mathbb{R}$. A basic difficulty of solving such problems is that the objective function $f(x)$ is given implicitly as the expectation and as such is difficult to compute to high accuracy. A way of solving problems (1.1) is by using randomized algorithms, based on Monte Carlo sampling. There are two competing approaches of this type, namely, the Sample Average Approximation (SAA) and the Stochastic Approximation (SA) methods. Both approaches have a long history.

The basic idea of the SAA method is to generate a sample ξ_1, \dots, ξ_N , of N realizations of ξ and to approximate the “true” problem (1.1) by replacing $f(x)$ with its sample average approximation $\hat{f}_N(x) := N^{-1} \sum_{t=1}^N F(x, \xi_t)$. Recent theoretical studies (cf., [2, 14, 15]) and numerical experiments (e.g., [4, 5, 16]) show that the SAA method coupled with a good deterministic algorithm for minimizing the constructed SAA problem could be reasonably efficient for solving certain classes of SP problems. The SA approach

Guanghui Lan

Georgia Institute of Technology, Atlanta, Georgia 30332, USA, E-mail: glan@isye.gatech.edu.

Arkadi Nemirovski

Georgia Institute of Technology, Atlanta, Georgia 30332, USA, E-mail: nemirovs@isye.gatech.edu, research of this author was partly supported by the NSF award DMI-0619977.

Alexander Shapiro

Georgia Institute of Technology, Atlanta, Georgia 30332, USA, E-mail: ashapiro@isye.gatech.edu, research of this author was partly supported by the NSF award DMI-0619977.

originates from the pioneering work of Robbins and Monro [12] and was discussed in numerous publications since. An important improvement was developed in Polyak [10] and Polyak and Juditsky [11], where a robust version of the SA method was introduced (the main ingredients of Polyak's scheme, long steps and averaging, were in a different form proposed already in Nemirovski and Yudin [6]). Yet it was believed that the SA approach performs poorly in practice and cannot compete with the SAA method. Somewhat surprisingly it was demonstrated recently in Nemirovski et al. [8] that a proper modification of the SA approach, based on the Nemirovski and Yudin [7] mirror-descent method, can be competitive and can even significantly outperform the SAA method for a certain class of convex stochastic programs. For example, when X in (1.1) is a simplex of large dimension, the *Mirror Descent Stochastic Approximation* builds approximate solutions 10 – 40 times faster than an SAA based algorithm while keeping similar solution quality.

An important methodological property of the SAA approach is that, with some additional effort, it can provide an estimate of the accuracy of an obtained solution by computing upper and lower (confidence) bounds for the optimal value of the true problem (cf., [5,9]). The main goal of this paper is to show that, for a certain class of stochastic convex problems, the Mirror Descent SA method can also provide similar bounds with considerably less computational effort. More specifically we study in this paper the following aspects of the Mirror Descent SA method.

- Investigate different ways to estimate lower and upper bounds for the objective values by the Mirror Descent SA method, and thus to obtain an accuracy certificate for the attained solutions.
- Adjust the Mirror Descent SA method to solve two interesting application problems in asset allocation, namely, minimizing¹ the Expected Utility (EU) and minimizing the Conditional Value-at-Risk (CVaR). These models are widely used in practice, for example, by investment companies, brokerage firms, mutual funds, and any business that evaluates risks (cf., [13]).
- Understand the performance of the Mirror Descent SA algorithm for solving stochastic programs with a feasible region more complicated than a simplex. For the EU model, the feasible region is the intersection of a simplex with a box constraint and we will compare two different variants of SA methods for solving it. For the CVaR problem, the feasible region is a polyhedron and we will discuss some techniques to explore its structure.

The paper is organized as follows. In section 2 we briefly introduce the Mirror Descent SA method. Section 3 is devoted to a derivation and analysis of statistical upper and lower bounds for the optimal value of the true problem. In section 4 we discuss an application of the Mirror Descent SA method to the expected utility and conditional value at risk approaches for the asset allocation problem. A discussion of numerical results is presented in section 5. Finally, proofs of technical results are given in the Appendix.

We assume throughout the paper that for every $\xi \in \Xi$ the function $F(\cdot, \xi)$ is convex on X , and that the expectation

$$\mathbb{E}[F(x, \xi)] = \int_{\Xi} F(x, \xi) dP(\xi) \tag{1.2}$$

is well defined, finite valued and continuous at every $x \in X$. That is, the expectation function $f(x)$ is finite valued, convex and continuous on X . For a norm $\|\cdot\|$ on \mathbb{R}^n , we denote by $\|x\|_* := \sup\{x^T y : \|y\| \leq 1\}$ the conjugate norm. By $\|x\|_p$ we denote the ℓ_p norm of vector $x \in \mathbb{R}^n$. In particular, $\|x\|_2 = \sqrt{x^T x}$ is the Euclidean norm of $x \in \mathbb{R}^n$. By $\Pi_X(x) := \arg \min_{y \in X} \|x - y\|_2$ we denote metric projection operator onto X . For the process ξ_1, ξ_2, \dots , we set $\xi^t := (\xi_1, \dots, \xi_t)$, and denote by $\mathbb{E}_{|t}$ or by $\mathbb{E}[\cdot | \xi^t]$ the conditional, ξ^t being given, expectation. For a number $a \in \mathbb{R}$ we denote $[a]_+ := \max\{a, 0\}$. By $\partial\phi(x)$ we denote the subdifferential of a convex function $\phi(x)$.

2 The Mirror Descent Stochastic Approximation Method

In this section, we give a brief introduction to the Mirror Descent SA algorithm as presented in [8]. We equip the embedding space \mathbb{R}^n , of the feasible domain X of (1.1), with a norm $\|\cdot\|$. We say that a function

¹ In order to have a convex rather than concave objective function, we deal here with minimization rather than maximization of the Expected Utility.

$\omega : X \rightarrow \mathbb{R}$ is a *distance generating function* with respect to the norm $\|\cdot\|$ and modulus $\alpha > 0$, if the following conditions hold: (i) ω is convex and continuous on X , (ii) the set

$$X^o := \{x \in X : \partial\omega(x) \neq \emptyset\} \quad (2.1)$$

is convex, and (iii) $\omega(\cdot)$ restricted to X^o is continuously differentiable and strongly convex with parameter α with respect to $\|\cdot\|$, i.e.,

$$(x' - x)^T (\nabla\omega(x') - \nabla\omega(x)) \geq \alpha \|x' - x\|^2, \quad \forall x', x \in X^o. \quad (2.2)$$

Note that the set X^o always contains relative interior of the set X .

With the distance generating function $\omega(\cdot)$ are associated the *prox-function*² $V : X^o \times X \rightarrow \mathbb{R}_+$ defined as

$$V(x, z) := \omega(z) - \omega(x) - \nabla\omega(x)^T(z - x), \quad (2.3)$$

the *prox-mapping* $P_x : \mathbb{R}^n \rightarrow X^o$ defined as

$$P_x(y) := \arg \min_{z \in X} \{y^T(z - x) + V(x, z)\}, \quad (2.4)$$

and the constant

$$D_{\omega, X} := \sqrt{\max_{x \in X} \omega(x) - \min_{x \in X} \omega(x)}. \quad (2.5)$$

Let x_1 be the minimizer of $\omega(\cdot)$ over X . This minimizer exists and is unique since X is convex and compact and $\omega(\cdot)$ is continuous and strictly convex on X . Observe that $x_1 \in X^o$, and since x_1 is the minimizer of $\omega(\cdot)$ it follows that $(x - x_1)^T \nabla\omega(x_1) \geq 0$ for all $x \in X$. Combined with the strong convexity of $\omega(\cdot)$ this implies that

$$\frac{1}{2}\alpha \|x - x_1\|^2 \leq V(x_1, x) \leq \omega(x) - \omega(x_1) \leq D_{\omega, X}^2, \quad \forall x \in X, \quad (2.6)$$

and hence

$$\|x - x_1\| \leq A_{\omega, X} := \sqrt{\frac{2}{\alpha}} D_{\omega, X}, \quad \forall x \in X. \quad (2.7)$$

Throughout the paper we assume existence of the following *stochastic oracle*.

- It is possible to generate an iid sample ξ_1, ξ_2, \dots , of realizations of random vector ξ , and we have access to a “black box” subroutine (a stochastic oracle): given $x \in X$ and a random realization $\xi \in \Xi$, the oracle returns the quantity $F(x, \xi)$ and a *stochastic subgradient* – a vector $\mathbf{G}(x, \xi)$ such that $\mathbf{g}(x) := \mathbb{E}[\mathbf{G}(x, \xi)]$ is well defined and is a subgradient of $f(\cdot)$ at x , i.e., $\mathbf{g}(x) \in \partial f(x)$.

We also make the following assumption.

(A1) There are positive constants Q and M_* such that for any $x \in X$:

$$\mathbb{E} \left[(F(x, \xi) - f(x))^2 \right] \leq Q^2, \quad (2.8)$$

$$\mathbb{E} [\|\mathbf{G}(x, \xi)\|_*^2] \leq M_*^2. \quad (2.9)$$

² It is also called Bregman distance [1].

It could be noted that $\mathbb{E}[(F(x, \xi) - f(x))^2]$ in (2.8) is variance of the random variable $F(x, \xi)$.

When speaking about Stochastic Approximation as applied to minimization problem (1.1), one usually does not care of how the values of $f(\cdot)$ are observed. All what matters is the observations of the gradient, this is the only information used by the basic SA algorithm (2.10), see below. We, however, are interested in building upper and lower bounds on the optimal value and/or value of $f(\cdot)$ at a given solution, and in this respect, it does matter how these values are observed. Conditions (2.8)–(2.9) of assumption (A1) impose restrictions on the magnitudes of noises in the unbiased observations of the values of $f(\cdot)$ and the subgradients of $f(\cdot)$ reported by the stochastic oracle.

The description of the Mirror Descent SA algorithm is as follows. Starting from point x_1 , the algorithm iteratively generates points $x_t \in X^o$ according to the recurrence

$$x_{t+1} := P_{x_t}(\gamma_t \mathbf{G}(x_t, \xi_t)), \quad (2.10)$$

where $\gamma_t > 0$ are deterministic stepsizes. Note that for $\omega(x) := \frac{1}{2}\|x\|_2^2$, we have that $P_x(y) = \Pi_X(x - y)$ and hence $x_{t+1} = \Pi_X(x_t - \gamma_t \mathbf{G}(x_t, \xi_t))$. In that case, the Mirror Descent SA method is referred to as the *Euclidean SA*.

Now let N be the total number of steps. Let us set

$$\nu_t := \frac{\gamma_t}{\sum_{i=1}^N \gamma_i}, \quad t = 1, \dots, N, \quad \text{and} \quad \tilde{x}_N := \sum_{t=1}^N \nu_t x_t. \quad (2.11)$$

Note that $\sum_{t=1}^N \nu_t = 1$, and hence \tilde{x}_N is a convex combination of the iterates x_1, \dots, x_N . Here \tilde{x}_N is considered as the approximate solution generated by the algorithm in course of N steps. The quality of this solution can be quantified as follows (cf., [8, p.1583]).

Proposition 1 *Suppose that condition (2.9) of assumption (A1) holds. Then for the N -step of Mirror Descent SA algorithm we have that*

$$\mathbb{E}[f(\tilde{x}_N) - \text{Opt}] \leq \frac{D_{\omega, X}^2 + (2\alpha)^{-1} M_*^2 \sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t}. \quad (2.12)$$

In implementations of the SA algorithm different stepsize strategies can be applied to (2.10) (see [8]). We discuss now the *constant stepsize* policy. That is, we assume that the number N of iterations is fixed in advance, and $\gamma_t = \gamma$, $t = 1, \dots, N$. In that case

$$\tilde{x}_N = \frac{1}{N} \sum_{t=1}^N x_t. \quad (2.13)$$

By choosing the stepsizes as

$$\gamma_t = \gamma := \frac{\theta \sqrt{2\alpha} D_{\omega, X}}{M_* \sqrt{N}}, \quad t = 1, \dots, N, \quad (2.14)$$

with a (scaling) constant $\theta > 0$, we have in view of (2.12) that

$$\mathbb{E}[f(\tilde{x}_N) - \text{Opt}] \leq \max\{\theta, \theta^{-1}\} \Lambda_{\omega, X} M_* N^{-1/2}, \quad (2.15)$$

with $\Lambda_{\omega, X}$ given by (2.7). This shows that scaling the stepsizes by the (positive) constant θ results in updating the estimate (2.15) by the factor of $\max\{\theta, \theta^{-1}\}$ at most. By Markov inequality it follows from (2.15) that for any $\varepsilon > 0$,

$$\text{Prob}\{f(\tilde{x}_N) - \text{Opt} > \varepsilon\} \leq \frac{\sqrt{2} \max\{\theta, \theta^{-1}\} D_{\omega, X} M_*}{\varepsilon \sqrt{\alpha N}}. \quad (2.16)$$

It is possible to obtain finer bounds for the probabilities in the left hand side of (2.16) when imposing conditions more restrictive than conditions of assumption(A1). Consider the following conditions.

(A2) There are positive constants Q and M_* such that for any $x \in X$:

$$\mathbb{E} \left[\exp \left\{ |F(x, \xi) - f(x)|^2 / Q^2 \right\} \right] \leq \exp\{1\}, \quad (2.17)$$

$$\mathbb{E} \left[\exp \left\{ \|\mathbf{G}(x, \xi)\|_*^2 / M_*^2 \right\} \right] \leq \exp\{1\}. \quad (2.18)$$

Note that conditions (2.17)–(2.18) are stronger than the respective conditions (2.8)–(2.9). Indeed, if a random variable Y satisfies $\mathbb{E}[\exp\{Y/a\}] \leq \exp\{1\}$ for some $a > 0$, then by Jensen inequality $\exp\{\mathbb{E}[Y/a]\} \leq \mathbb{E}[\exp\{Y/a\}] \leq \exp\{1\}$, and therefore $\mathbb{E}[Y] \leq a$. Of course, conditions (2.17)–(2.18) hold if for all $(x, \xi) \in X \times \Xi$:

$$|F(x, \xi) - f(x)| \leq Q \quad \text{and} \quad \|\mathbf{G}(x, \xi)\|_* \leq M_*.$$

The following result has been established in [8, Proposition 2.2].

Proposition 2 *Suppose that condition (2.18) of assumption (A2) holds. Then for the constant stepsize policy, with the stepsize (2.14), the following inequality holds for any $\Omega \geq 1$:*

$$\text{Prob} \left\{ f(\tilde{x}_N) - \text{Opt} > \max\{\theta, \theta^{-1}\} (12 + 2\Omega) A_{\omega, X} M_* N^{-1/2} \right\} \leq 2 \exp\{-\Omega\}. \quad (2.19)$$

It follows from (2.19) that the number N of steps required by the algorithm to solve the problem with accuracy $\varepsilon > 0$, and a (probabilistic) confidence $1 - \beta$, is of order $O\left(\varepsilon^{-2} \log^2(1/\beta)\right)$. Note also that in practice one can modify the Mirror Descent SA algorithm so that the approximate solution \tilde{x}_N is obtained by averaging over a part of the trajectory (see [8] for details).

3 Accuracy certificates for SA solutions

In this section, we discuss several ways to estimate lower and upper bounds for the optimal value of problem (1.1), which gives us an accuracy certificate for obtained solutions. Specifically, we distinguish between two types of certificates: the *online certificates* that can be computed quickly when running the SA algorithm, and the *offline certificates* obtained in a more time consuming way at the dedicated *validation step*, after a solution has been obtained.

3.1 Online certificate

Consider the numbers ν_t and solution \tilde{x}_N , defined in (2.11), functions

$$f^N(x) := \sum_{t=1}^N \nu_t \left[f(x_t) + \mathbf{g}(x_t)^T (x - x_t) \right] \quad \text{and} \quad \hat{f}^N(x) := \sum_{t=1}^N \nu_t [F(x_t, \xi_t) + \mathbf{G}(x_t, \xi_t)^T (x - x_t)],$$

and define

$$f_*^N := \min_{x \in X} f^N(x) \quad \text{and} \quad f^{*N} := \sum_{t=1}^N \nu_t f(x_t). \quad (3.1)$$

Since $\nu_t > 0$ and $\sum_{t=1}^N \nu_t = 1$, it follows by convexity of $f(\cdot)$ that the function $f^N(\cdot)$ underestimates $f(\cdot)$ everywhere on X , and hence $f_*^N \leq \text{Opt}$. Since $\tilde{x}_N \in X$ we also have that $\text{Opt} \leq f(\tilde{x}_N)$, and by convexity of $f(\cdot)$ that $f(\tilde{x}_N) \leq f^{*N}$. That is, for *any realization* of the random sample ξ_1, \dots, ξ_N we have that

$$f_*^N \leq \text{Opt} \leq f(\tilde{x}_N) \leq f^{*N}. \quad (3.2)$$

It follows from (3.2) that $\mathbb{E}[f_*^N] \leq \text{Opt} \leq \mathbb{E}[f^{*N}]$ as well.

Of course, the bounds f_*^N and f^{*N} are unobservable since the values $f(x_t)$ are not known exactly. Therefore we consider their computable counterparts

$$\underline{f}^N = \min_{x \in X} \hat{f}^N(x) \quad \text{and} \quad \bar{f}^N = \sum_{t=1}^N \nu_t F(x_t, \xi_t). \quad (3.3)$$

We refer to \underline{f}^N and \bar{f}^N as *online* bounds. The bound \bar{f}^N can be easily calculated while running the SA procedure. The bound \underline{f}^N involves solving the optimization problem of minimizing a linear in x objective function over set X . If the set X is defined by linear constraints, this is a linear programming problem.

Since x_t is a function of $\xi^{t-1} = (\xi_1, \dots, \xi_{t-1})$, and ξ_t is independent of ξ^{t-1} , we have that

$$\mathbb{E}[\bar{f}^N] = \sum_{t=1}^N \nu_t \mathbb{E}\{\mathbb{E}[F(x_t, \xi_t) | \xi^{t-1}]\} = \sum_{t=1}^N \nu_t \mathbb{E}[f(x_t)] = \mathbb{E}[f^{*N}]$$

and

$$\begin{aligned} \mathbb{E}[\underline{f}^N] &= \mathbb{E}\left[\mathbb{E}\left\{\min_{x \in X} \left[\sum_{t=1}^N \nu_t [F(x_t, \xi_t) + \mathbf{G}(x_t, \xi_t)^T(x - x_t)]\right] \middle| \xi^{t-1}\right\}\right] \\ &\leq \mathbb{E}\left[\min_{x \in X} \left\{\mathbb{E}\left[\sum_{t=1}^N \nu_t [F(x_t, \xi_t) + \mathbf{G}(x_t, \xi_t)^T(x - x_t)]\right] \middle| \xi^{t-1}\right\}\right] \\ &= \mathbb{E}\left[\min_{x \in X} f^N(x)\right] = \mathbb{E}[f_*^N]. \end{aligned}$$

It follows that

$$\mathbb{E}[\underline{f}^N] \leq \text{Opt} \leq \mathbb{E}[\bar{f}^N]. \quad (3.4)$$

That is, on average \underline{f}^N and \bar{f}^N give, respectively, a lower and an upper bound for the optimal value of problem (1.1). In order to see how good are the bounds \underline{f}^N and \bar{f}^N let us estimate expectations and probabilities of the corresponding errors. Proof of the following theorem is given in the Appendix.

Theorem 1 (i) *Suppose that assumption (A1) holds. Then*

$$\mathbb{E}[f^{*N} - f_*^N] \leq \frac{2D_{\omega, X}^2 + \frac{5}{2}\alpha^{-1}M_*^2 \sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t}, \quad (3.5)$$

$$\mathbb{E}\left[|\bar{f}^N - f^{*N}|\right] \leq Q \sqrt{\sum_{t=1}^N \nu_t^2}, \quad (3.6)$$

$$\mathbb{E}\left[|\underline{f}^N - f_*^N|\right] \leq \frac{D_{\omega, X}^2 + \frac{1}{2}\alpha^{-1}M_*^2 \sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t} + (Q + 8A_{\omega, X}M_*) \sqrt{\sum_{t=1}^N \nu_t^2}. \quad (3.7)$$

In particular, in the case of constant stepsize policy (2.14) we have

$$\begin{aligned} \mathbb{E}[f^{*N} - f_*^N] &\leq \left[\theta^{-1} + 5\theta/2\right] A_{\omega, X} M_* N^{-1/2}, \\ \mathbb{E}\left[|\bar{f}^N - f^{*N}|\right] &\leq Q N^{-1/2}, \\ \mathbb{E}\left[|\underline{f}^N - f_*^N|\right] &\leq \frac{1}{2} \left[\theta^{-1} + \theta\right] A_{\omega, X} M_* N^{-1/2} + (Q + 8A_{\omega, X}M_*) N^{-1/2}, \end{aligned} \quad (3.8)$$

where $A_{\omega, X}$ is given by (2.7).

(ii) *Moreover, if assumption (A1) is strengthened to assumption (A2), then in the case of constant³ stepsize policy (2.14) we have for any $\Omega \geq 0$:*

³ The bounds in the Appendix cover the case of general-type stepsizes; here we restrict ourselves with the case of constant stepsizes to avoid less transparent formulas.

$$\text{Prob} \left\{ f^{*N} - f_*^N > N^{-1/2} \Lambda_{\omega, X} M_* \left(\left[\frac{5}{2} \theta + \theta^{-1} \right] + \Omega \left[4 + \frac{5}{2} \theta N^{-1/2} \right] \right) \right\} \leq 2 \exp\{-\Omega^2/3\} + 2 \exp\{-\Omega^2/12\} + 2 \exp\{-3\Omega\sqrt{N}/4\}, \quad (3.9)$$

$$\text{Prob} \left\{ |\bar{f}^N - f^{*N}| > \Omega Q \sqrt{\sum_{t=1}^N \nu_t^2} \right\} \leq 2 \exp\{-\Omega^2/3\}, \quad (3.10)$$

$$\text{Prob} \left\{ |\underline{f}^N - f_*^N| > N^{-1/2} \left(\left[\frac{1}{2\theta} + 2\theta \right] \Lambda_{\omega, X} M_* + \Omega \left[Q + [8 + 2\theta N^{-1/2}] \Lambda_{\omega, X} M_* \right] \right) \right\} \leq 6 \exp\{-\Omega^2/3\} + \exp\{-\Omega^2/12\} + \exp\{-3\Omega\sqrt{N}/4\}. \quad (3.11)$$

Estimates of the above theorem show that as N grows, the observable quantities \underline{f}^N and \bar{f}^N approach, in a probabilistic sense, their unobservable counterparts, which, in turn, approach each other and thus the optimal value of problem (1.1). For the constant stepsize policy (2.14), we have that all estimates given in the right hand side of (3.8) are of order $O(N^{-1/2})$. It follows that under assumption (A1) and for the constant stepsize policy, difference between the upper \bar{f}^N and lower \underline{f}^N bounds converges on average to zero, with increase of the sample size N , at a rate of $O(N^{-1/2})$.

Note that for the constant stepsize policy (2.14) and under assumption (A2), the bounds (3.9) – (3.11) combine with (3.2) to imply that

- $\text{Prob} \left\{ \bar{f}_\Omega^N := \bar{f}^N + \Omega \sigma_+ N^{-1/2} \text{ is not an upper bound on } f(\tilde{x}^N) \right\} \leq 2e^{-\frac{\Omega^2}{3}}$, with $\sigma_+ = Q$;
- $\text{Prob} \left\{ \underline{f}_\Omega^N := \underline{f}^N - [\mu_- + \Omega \sigma_-] N^{-1/2} \text{ is not a lower bound on Opt} \right\} \leq 6e^{-\frac{\Omega^2}{3}} + e^{-\frac{\Omega^2}{12}} + e^{-\frac{3\Omega\sqrt{N}}{4}}$, with $\Lambda_{\omega, X}$ defined by (2.7) and

$$\mu_- := \left[\frac{1}{2\theta} + 2\theta \right] \Lambda_{\omega, X} M_*, \quad \sigma_- := Q + [8 + 2\theta N^{-1/2}] \Lambda_{\omega, X} M_*;$$

- $\text{Prob} \left\{ \bar{f}_\Omega^N - \underline{f}_\Omega^N > [\mu + \Omega \sigma] N^{-1/2} \right\} \leq 10e^{-\frac{\Omega^2}{3}} + 3e^{-\frac{\Omega^2}{12}} + 3e^{-\frac{3\Omega\sqrt{N}}{4}}$, with

$$\mu := \left[\frac{3}{2\theta} + \frac{9\theta}{2} \right] \Lambda_{\omega, X} M_*, \quad \sigma := 2Q + \left[12 + \frac{9\theta}{2} \right] \Lambda_{\omega, X} M_*.$$

Theorem 1 shows that for large N the online observable random quantities \bar{f}^N and \underline{f}^N are close to the upper bound f^{*N} and lower bound f_*^N , respectively. Besides this, on average, \bar{f}^N indeed overestimates Opt, and \underline{f}^N indeed underestimates Opt. To save words, let us call random estimates which on average under- or overestimate a certain quantity, *on average* lower, respectively, upper bounds on this quantity. From now on, when speaking of “true” lower and upper bounds – those which always (or almost surely) under-, respectively, overestimate the quantity, we add the adjective “valid”. Thus, we refer to f^{*N} and f_*^N as *valid* upper and lower bounds on Opt, respectively. Recall that f^{*N} is also a valid upper bound on $f(\tilde{x}_N)$.

Remark 1 Recall that the SAA approach also provides a lower on average bound – the random quantity \hat{f}_{SAA}^N , which is the optimal value of the sample average problem (cf., [5, 9]). Suppose the same sample ξ_t , $t = 1, \dots, N$, is applied for both SA and SAA methods. Besides this, assume that the constant stepsize policy is used in the SA method, and hence $\nu_t = 1/N$, $t = 1, \dots, N$. Finally, assume (as it often is the case) that $\mathbf{G}(x, \xi)$ is a subgradient of $F(x, \xi)$ in x . By convexity of $F(\cdot, \xi)$ and since $\underline{f}^N = \min_{x \in X} \hat{f}^N(x)$, we have

$$\hat{f}_{\text{SAA}}^N := \min_{x \in X} N^{-1} \sum_{t=1}^N F(x, \xi_t) \geq \min_{x \in X} \sum_{t=1}^N \nu_t \left(F(x_t, \xi_t) + \mathbf{G}(x_t, \xi_t)^T (x - x_t) \right) = \underline{f}^N. \quad (3.12)$$

That is, for the same sample the lower bound \underline{f}^N is smaller than the lower bound obtained by the SAA method. However, it should be noted that the lower bound \underline{f}^N is computed much faster than \hat{f}_{SAA}^N , since

computing the latter one amounts to solving the sample average optimization problem associated with the generated sample. Moreover, we will discuss in the next subsection how to improve the lower bound \underline{f}^N . From the computational results, the improved lower bound is comparable to the one obtained by the SAA method. ■

Remark 2 Similar to the SAA method, in order to estimate the variability of the lower bound \underline{f}^N , one can run the SA procedure M times, with independent samples, each of size N , and consequently compute the average and sample variance of M realizations of the random quantity \underline{f}^N . Alternatively, one can run the SA procedure once but with NM iterations, then partition the obtained trajectory into M consecutive parts, each of size N , for each of these parts calculate the corresponding SA lower bound and consequently compute the average and sample variance of the M obtained numbers. ■

3.2 Offline certificate

Suppose now that the Mirror Descent SA method is terminated after N iterations. Given a solution \tilde{x}_N obtained by this method, the objective value $f(\tilde{x}_N)$ can be estimated by Monte Carlo sampling. That is, an iid random sample ξ_j , $j = 1, \dots, K$, (independent of the random sample used in computing \tilde{x}_N) is generated and $f(\tilde{x}_N)$ is estimated by $\hat{f}_K(\tilde{x}_N) := K^{-1} \sum_{j=1}^K F(\tilde{x}_N, \xi_j)$. Since this procedure does not require computing prox-mapping and the like, one can use here a large sample size K . Of course, we can expect that $\hat{f}_K(\tilde{x}_N)$ is a better upper bound on $f(\tilde{x}_N)$ than the online counterpart \bar{f}^N of the valid upper bound f^{*N} .

We now demonstrate that the online lower bound \underline{f}^N can be also improved in the validation step. Given an iid random sample ξ_j , $j = 1, \dots, L$, we can estimate the (linear in x) form $\ell_L(x; \tilde{x}_N) := f(\tilde{x}_N) + \mathbf{g}(\tilde{x}_N)^T(x - \tilde{x}_N)$ by

$$\hat{\ell}_L(x; \tilde{x}_N) := \frac{1}{L} \sum_{j=1}^L \left[F(\tilde{x}_N, \xi_j) + \mathbf{G}(\tilde{x}_N, \xi_j)^T(x - \tilde{x}_N) \right], \quad (3.13)$$

and hence construct the following lower bound of Opt:

$$\underline{lb}^N := \min_{x \in X} \left\{ \max \left[\hat{f}^N(x), \hat{\ell}_L(x; \tilde{x}_N) \right] \right\}. \quad (3.14)$$

Clearly, by definition we have that $\underline{lb}^N \geq \underline{f}^N$.

It is worth of noting that although $\mathbb{E}[\hat{f}^N(x)] \leq f(x)$ and $\mathbb{E}[\hat{\ell}_L(x; \tilde{x}_N)] \leq f(x)$, the expected value of the maximum of these two quantities is not necessarily $\leq f(x)$. Therefore the expected value of \underline{lb}^N is not necessarily $\leq \text{Opt}$, i.e., we cannot claim that \underline{lb}^N is a lower on average bound on Opt. However, the following results shows that \underline{lb}^N is “statistically close” to a valid lower bound on Opt, provided that N and L are large. Proof of the following theorem is given in the Appendix.

Theorem 2 *Suppose that assumption (A1) holds and let the constant stepsizes (2.14) be used. Then*

$$\sqrt{\mathbb{E} \left\{ \left(\left[\underline{lb}^N - \text{Opt} \right]_+ \right)^2 \right\}} \leq \sqrt{2Q^2 + 32A_{\omega, X}^2 M_*^2} \left[\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{L}} \right]. \quad (3.15)$$

Moreover, under assumption (A2), we have that for all $\Omega \geq 0$:

$$\text{Prob} \left\{ \underline{lb}^N - \text{Opt} > \left[Q + 4A_{\omega, X} M_* \right] \left[\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{L}} \right] \right\} \leq 4 \exp\{-\Omega^2/3\}. \quad (3.16)$$

4 Applications in Asset Allocation

In this section, we discuss an application of the Mirror Descent SA method to solving asset allocation problems based on the Expected Utility (EU) and the Conditional Value-at-Risk (CVaR) models.

4.1 Minimizing the expected utility

We consider the following stochastic utility model:

$$\min_{x \in X} \{f(x) := \mathbb{E}[\phi(\sum_{i=1}^n (a_i + \xi_i)x_i)]\}. \quad (4.1)$$

Here $X := X' \cap X''$, where

$$X' := \{x \in \mathbb{R}^n : \sum_{i=1}^n x_i \leq r\} \quad \text{and} \quad X'' := \{x \in \mathbb{R}^n : l_i \leq x_i \leq u_i, i = 1, \dots, n\},$$

$r > 0$, a_i and $0 \leq l_i < u_i$, $i = 1, \dots, n$, are given numbers, $\xi_i \sim \mathcal{N}(0, 1)$ are independent random variables having standard normal distribution and $\phi(\cdot)$ is a piecewise linear convex function given by

$$\phi(t) := \max\{c_1 + b_1 t, \dots, c_m + b_m t\}, \quad (4.2)$$

where c_j and b_j , $j = 1, \dots, m$, are certain constants. Note that by varying parameters r and l_i, u_i we can change the feasible region from a simplex to a box, or the intersection of a simplex with a box. Note that since the set X is compact and $f(x)$ is continuous, the set of optimal solution of (4.1) is nonempty, provided that X is nonempty. A simpler version of problem (4.1), in which X is assumed to be a standard simplex, has been considered in [8].

For solving this problem, we consider two variants of the Mirror Descent SA algorithm: *Non-Euclidean SA (N-SA)* and *Euclidean SA (E-SA)*, which differ from each other in how the norm $\|\cdot\|$ and the distance generating function $\omega(\cdot)$ are chosen.

4.1.1 Non-Euclidean SA

In N-SA for solving the EU model, the entropy distance generating function

$$\omega(x) := \sum_{i=1}^n \frac{x_i}{r} \ln \frac{x_i}{r}, \quad (4.3)$$

coupled with the $\|\cdot\|_1$ norm is employed. Note that here $X^o = \{x \in X : x > 0\}$ and for $n \geq 3$,

$$D_{\omega, X}^2 = \max_{x \in X} \omega(x) - \min_{x \in X} \omega(x) \leq \max_{x \in X'} \omega(x) - \min_{x \in X'} \omega(x) \leq \ln n.$$

Also observe that for any $x \in X'$, $x > 0$, and $h \in \mathbb{R}^n$,

$$\begin{aligned} \left(\sum_{i=1}^n |h_i|\right)^2 &= \left(\sum_{i=1}^n x_i^{1/2} |h_i| x_i^{-1/2}\right)^2 \leq \left(\sum_{i=1}^n x_i\right) \left(\sum_{i=1}^n h_i^2 x_i^{-1}\right) \\ &\leq r \left(\sum_{i=1}^n h_i^2 x_i^{-1}\right) = r^2 h^T \nabla^2 \omega(x) h, \end{aligned}$$

where the first inequality follows by Cauchy inequality. Therefore the modulus of ω , with respect to the $\|\cdot\|_1$ norm, satisfies $\alpha \geq r^{-2}$. Note that here $D_{\omega, X}$ can be overestimated while α being underestimated since $X' \subseteq X$, therefore, the stepsizes computed according to (2.14) in view of these estimates may not be optimal. Of course, the quantity $D_{\omega, X}$ can be estimated more accurately, for example, by computing $\min_{x \in X} \omega(x)$ explicitly. We will also discuss a few different ways to fine-tune the stepsizes in Section 5.

For the entropy distance generating function (4.3), the prox-mapping $P_v(z)$ (defined in (2.4)) is r times the optimal solution to the optimization problem

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n (s_i x_i + x_i \ln x_i), \\ \text{s.t.} \quad & \sum_{i=1}^n x_i \leq 1, \\ & \tilde{l}_i \leq x_i \leq \tilde{u}_i, \quad i = 1, \dots, n, \end{aligned} \quad (4.4)$$

where $s_i = rz_i - \ln(v_i/r) - 1$, $\tilde{l}_i = l_i/r$, $\tilde{u}_i = u_i/r$.

In some cases problem (4.4) has an explicit solution, e.g., if $l_i = 0$ and $u_i \geq r$, $i = 1, \dots, n$ (in that case the constraints $z_i \leq u_i$ are redundant). In general, we can solve (4.4) as follows. Let $\lambda \geq 0$ denote the

Lagrange multiplier associated with the constraint $\sum_{i=1}^n x_i \leq 1$ and consider the corresponding Lagrangian relaxation of (4.4):

$$\begin{aligned} \min_x \quad & \sum_{i=1}^n (s_i x_i + x_i \ln x_i) + \lambda \left(\sum_{i=1}^n x_i \right), \\ \text{s.t.} \quad & \bar{l}_i \leq x_i \leq \bar{u}_i, \quad i = 1, \dots, n. \end{aligned} \quad (4.5)$$

This is a separable problem. Since $s_i x_i + x_i \ln x_i + \lambda x_i$ is monotonically decreasing for x_i less than $\exp[-(s_i + 1 + \lambda)]$ and is monotonically increasing after, we have that the i -th coordinate $\bar{x}_i(\lambda)$ of the optimal solution of (4.5) is given by the projection of $\exp[-(s_i + 1 + \lambda)]$ onto the interval $[\bar{l}_i, \bar{u}_i]$. Then, to solve problem (4.4) is equivalent to find $\lambda \geq 0$ such that

$$\sum_{i=1}^n \bar{x}_i(\lambda) = 1, \quad \text{if } \lambda > 0, \quad (4.6)$$

$$\sum_{i=1}^n \bar{x}_i(\lambda) \leq 1, \quad \text{if } \lambda = 0. \quad (4.7)$$

While inequality (4.7) can be easily checked, the root-finding problem (4.6) is usually solved to certain precision by using bisection, and each bisection step requires $\mathcal{O}(n)$ operations.

4.1.2 Euclidean SA

In the E-SA approach to order to solve the EU model, the Euclidean distance generating function $\omega(x) := \frac{1}{2} x^T x$, coupled with the $\|\cdot\|_2$ norm is employed. Clearly here $X^o = X$ and $\alpha = 1$. We have

$$D_{\omega, X}^2 = \max_{x \in X} \omega(x) - \min_{x \in X} \omega(x) \leq \frac{1}{2} \left(\min\{r^2, \|u\|_2^2\} - \|l\|_2^2 \right).$$

Moreover a procedure similar to the one given in Subsection 4.1.1 can be developed for computing the prox mapping $P_x(y)$, which is given here by the metric projection $\Pi_X(x - y)$.

As it was noted in [8, Example 2.1], if X is a standard simplex, N-SA can be potentially $\mathcal{O}(\sqrt{n/\log n})$ times faster than E-SA. The same conclusion seems to be applicable to our current situation, although certain caution should be taken since the error estimate (2.14) now also depends on l , u and r .

4.2 Minimizing the Conditional Value-at-Risk

The idea of minimizing CVaR in place of Value-at-Risk (VaR) is due to Rockafellar and Uryasev [13]. Recall that VaR and CVaR of a random variable Z are defined as

$$\text{VaR}_{1-\beta}(Z) := \inf \{ \tau : \text{Prob}(Z \leq \tau) \geq 1 - \beta \}, \quad (4.8)$$

$$\text{CVaR}_{1-\beta}(Z) := \inf_{\tau \in \mathbb{R}} \left\{ \tau + \beta^{-1} \mathbb{E}[Z - \tau]_+ \right\}. \quad (4.9)$$

Note that

$$\text{VaR}_{1-\beta}(Z) \in \text{Argmin}_{\tau \in \mathbb{R}} \left\{ \tau + \beta^{-1} \mathbb{E}[Z - \tau]_+ \right\}, \quad (4.10)$$

and hence $\text{VaR}_{1-\beta}(Z) \leq \text{CVaR}_{1-\beta}(Z)$.

The problem of interest in this subsection is:

$$\min_{y \in Y} \text{CVaR}_{1-\beta}(-\xi^T y), \quad (4.11)$$

where ξ is a random vector with mean $\bar{\xi} := \mathbb{E}[\xi]$ and covariance matrix Σ , and

$$Y := \left\{ y \in \mathbb{R}_+^n : \sum_{i=1}^n y_i = 1, \bar{\xi}^T y \geq R \right\}.$$

We assume that Y is nonempty and, moreover, contains a positive point. For simplicity we assume in the remaining part of the paper that ξ has continuous distribution, and hence $\xi^T y$ has continuous distribution for any $y \in Y$.

In view of the definition of CVaR in (4.9), our problem becomes:

$$\min_{x \in X} f(x) := \tau + \frac{1}{\beta} \mathbb{E} \left\{ [-\xi^T y - \tau]_+ \right\}, \quad (4.12)$$

where $X := Y \times \mathbb{R}$ and $x := (y, \tau)$. Apparently, there exists one difficulty to apply the Mirror Descent SA for solving the above problem — in (4.12), the variables are y and τ , so that the feasible domain $Y \times \mathbb{R}$ of the problem is unbounded, while our Mirror Descent SA requires a bounded feasible domain. However, we will alleviate this problem by showing that the variable τ can actually be restricted into a bounded interval and thus the Mirror Descent SA method can be applied.

Noting that $\text{VaR}_{1-\beta}(Z) \in \text{Argmin}_{\tau \in \mathbb{R}} [\tau + \mathbb{E}\{[Z - \tau]_+\}]$, all we need is to find an interval which covers all points $\text{VaR}_{1-\beta}(-\xi^T y)$, $y \in Y$. Now, let Z be a random variable with finite mean μ and variance σ^2 . By Cantelli's inequality (also called one-sided Tschebyshev inequality) we have

$$\text{Prob}\{Z \geq t\} \leq \frac{\sigma^2}{(t - \mu)^2 + \sigma^2}.$$

Assuming that Z has continuous distribution, we obtain

$$\beta = \text{Prob}\{Z \geq \text{VaR}_{1-\beta}(Z)\} \leq \frac{\sigma^2}{[\text{VaR}_{1-\beta}(Z) - \mu]^2 + \sigma^2},$$

which implies that

$$\text{VaR}_{1-\beta}(Z) \leq \mu + \sqrt{\frac{1-\beta}{\beta}} \sigma. \quad (4.13)$$

Similarly, if $\text{VaR}_{1-\beta}(Z) \leq \mu$, then

$$1 - \beta = \text{Prob}\{-Z \geq -\text{VaR}_{1-\beta}(Z)\} \leq \frac{\sigma^2}{[-\text{VaR}_{1-\beta}(Z) + \mu]^2 + \sigma^2},$$

which implies that

$$\text{VaR}_{1-\beta}(Z) \geq \mu - \sqrt{\frac{\beta}{1-\beta}} \sigma. \quad (4.14)$$

Combing inequality (4.13) and (4.14) we obtain

$$\text{VaR}_{1-\beta}(Z) \in \left[\mu - \sqrt{\frac{\beta}{1-\beta}} \sigma, \mu + \sqrt{\frac{1-\beta}{\beta}} \sigma \right]. \quad (4.15)$$

Note also that if Z is symmetric and $\beta \leq 0.5$, then the previous inclusion can be strengthened to

$$\text{VaR}_{1-\beta}(Z) \in \left[\mu, \mu + \sqrt{\frac{1-\beta}{\beta}} \sigma \right]. \quad (4.16)$$

From this analysis it clearly follows that we lose nothing when restricting τ in (4.12) to vary in the segment

$$\tau \in \mathcal{T} := \left[\underline{\mu} - \sqrt{\frac{\beta}{1-\beta}} \bar{\sigma}, \bar{\mu} + \sqrt{\frac{1-\beta}{\beta}} \bar{\sigma} \right], \quad (4.17)$$

where

$$\underline{\mu} := \min_{y \in Y} \{-\xi^T y\}, \quad \bar{\mu} := \max_{y \in Y} \{-\xi^T y\}, \quad \bar{\sigma}^2 := \max_{y \in Y} y^T \Sigma y. \quad (4.18)$$

In the case when ξ is symmetric and $\beta \leq 0.5$, this segment can be further reduced to:

$$\tau \in \mathcal{T}' := \left[\underline{\mu}, \bar{\mu} + \sqrt{\frac{1-\beta}{\beta}} \bar{\sigma} \right]. \quad (4.19)$$

Note that the quantities $\underline{\mu}$ and $\bar{\mu}$ can be easily computed by solving the corresponding linear programs in (4.18). Moreover, although $\bar{\sigma}$ can be difficult to compute exactly, it can be replaced with its easily computable upper bound $\max_i \Sigma_{ii}$.

It is worth noting that an alternative upper bound for τ can be obtained in some cases: given an initial point $y_0 \in Y$, we have

$$\text{CVaR}_{1-\beta}(-\xi^T y_0) \geq \text{CVaR}_{1-\beta}(-\xi^T y^*) \geq \text{VaR}_{1-\beta}(-\xi^T y^*),$$

where y^* is an optimal solution of problem (4.11). Therefore, in view of (4.10), if the value of $\text{CVaR}_{1-\beta}(-\xi^T y_0)$ can be computed or estimated (e.g., by Monte-Carlo simulation), we can restrict the variable τ in (4.12) to be $\leq \text{CVaR}_{1-\beta}(-\xi^T y_0)$.

To apply the Mirror Descent SA to problem (4.11), we set $X = Y \times \mathcal{T}$ and define the stochastic oracle by setting

$$\begin{aligned} F(x, \xi) &\equiv F(y, \tau, \xi) = \tau + \frac{1}{\beta} \max[-\xi^T y - \tau, 0], \\ \mathbf{G}(x, \xi) &\equiv [\mathbf{G}_y(y, \tau, \xi); \mathbf{G}_\tau(y, \tau, \xi)] = \begin{cases} [-\beta^{-1}\xi; 1 - \beta^{-1}], & -\xi^T y - \tau > 0 \\ [0; \dots; 0; 1] & , \text{otherwise} \end{cases} \end{aligned}$$

Further, we choose D_y and D_τ from the relations

$$D_y \geq \max \left[1/2, \sqrt{\max_{y \in Y} \sum_i y_i \ln y_i - \min_{y \in Y} \sum_i y_i \ln y_i} \right], \quad D_\tau = \frac{1}{2} \left[\max_{\tau \in \mathcal{T}} \tau^2 - \min_{\tau \in \mathcal{T}} \tau^2 \right]$$

(we always can take $D_y = \max[1/2, \sqrt{\ln(n)}]$) and equip X and its embedding space $\mathbb{R}_y^n \times \mathbb{R}_\tau \supset X$ with the distance generating function and the norm as follows:

$$\begin{aligned} \|(y, \tau)\| &= \sqrt{\|y\|_1^2 / (2D_y^2) + \tau^2 / (2D_\tau^2)} \quad [\Leftrightarrow \|(z, \rho)\|_* = \sqrt{2D_y^2 \|z\|_\infty^2 + 2D_\tau^2 \rho^2}] \\ \omega(x) &\equiv \omega(y, \tau) = \frac{1}{2D_y^2} \sum_{i=1}^n y_i \ln y_i + \frac{1}{2D_\tau^2} \tau^2 \end{aligned}$$

Note that with this setup, $X^o = \{(y, \tau) \in X : y > 0\}$. Besides this, it is easily seen that $\sum_{i=1}^n y_i \ln y_i$, restricted on Y , is strongly convex, modulus 1, w.r.t. $\|\cdot\|_1$, whence ω is strongly convex, modulus $\alpha = 1$, on X . An immediate computation shows that $D_{\omega, X} = 1$, and therefore $A_{\omega, X} = \sqrt{2}$. Finally, we set

$$M_* = \sqrt{2D_y^2 \beta^{-2} \mathbb{E}[\|\xi\|_\infty^2] + 2D_\tau^2 \max[1, (\beta^{-1} - 1)^2]}. \quad (4.20)$$

It is easy to verify that with this M_* , our stochastic oracle satisfies (2.9).

Indeed, from the formula for $\mathbf{G}(x, \xi)$ we have

$$\mathbb{E}[\|\mathbf{G}(x, \xi)\|_*^2] = \mathbb{E}[2D_y^2 \beta^{-2} \|\xi\|_\infty^2 + 2D_\tau^2 \max[1, \beta^{-1} - 1]^2] = M_*^2,$$

as required in (2.9). Further, for $x \in X$ we have $|F(x, \xi) - \tau - \beta^{-1} \max[-\tau, 0]| \leq \beta^{-1} |\xi^T y| \leq \beta^{-1} \|\xi\|_\infty$, whence

$$\begin{aligned} \mathbb{E}[(F(x, \xi) - f(x))^2] &= \mathbb{E}[(F(x, \xi) - \mathbb{E}[F(x, \xi)])^2] \leq \mathbb{E}[(F(x, \xi) - \tau - \beta^{-1} \max[-\tau, 0])^2] \\ &\leq \beta^{-2} \mathbb{E}[\|\xi\|_\infty^2] \leq A_{\omega, X}^2 M_*^2, \end{aligned}$$

where the concluding inequality is due to $D_y \geq 1/2$ and $A_{\omega, X} = \sqrt{2}$. We see that assumption (A1) is satisfied with M_* given by (4.20) and $Q = A_{\omega, X} M_* = \sqrt{2} M_*$.

Table 1 The test instances for EU model

NAME	r	u	NAME	r	u
EU-1	100	0.05	EU-6	1	$+\infty$
EU-2	100	0.20	EU-7	10	$+\infty$
EU-3	100	0.40	EU-8	100	$+\infty$
EU-4	100	10.00	EU-9	1,000	$+\infty$
EU-5	100	50.00	EU-10	5,000	$+\infty$

Table 2 The stepsize factors

NAME	Best θ	Inferred θ	NAME	Best θ	Inferred θ
EU-1	0.005	0.005	EU-6	5.000	5.000
EU-2	1.000	5.000	EU-7	10.000	10.000
EU-3	1.000	5.000	EU-8	10.000	10.000
EU-4	5.000	10.000	EU-9	10.000	10.000
EU-5	5.000	5.000	EU-10	5.000	5.000

5 Numerical results

5.1 More implementation details

- **Fine-tuning the stepsizes:** In Section 2, we specified the constant stepsize policy for the Mirror Descent SA method up to the “scaling parameter” θ . In our experiments, this parameter was chosen by as a result of pilot runs of the Mirror Descent SA algorithm with several trial values of θ and a very small sample size N (namely, $N = 100$). From these values of θ , we chose for the actual run the one resulting in the smallest online upper bound \bar{f}^N on the optimal value.
- **Bundle-level method for solving SAA problem:** We also compare the results obtained by the Mirror Descent SA method with those obtained by the SAA coupled with the bundle-level method (SAA-BL) [3]. Note that the SAA problem is to be solved by the Bundle-level method; in our experiments, the SAA problems were solved within relative accuracy 1.e-4 through 1.e-6, depending on the instance.

5.2 Computational results for the EU model

In our experiments, we fix $l_i = 0$ and $u_i = u$ for all $1 \leq i \leq n$. The experiments were conducted for ten random instances which have the same dimension $n = 1000$ but differ in the parameters u and r , and the function $\phi(\cdot)$. A detailed description of these instances is shown in Table 1. Observe that for the first five instances, we fix $r = 100$ but change u from 0.05 to 50. For the next five instances, we assume $u = +\infty$ but change r from 1.0 to 5,000.0.

Here we highlight some interesting findings based on our computational results. More numerical results can be found the end of this paper.

- **The effect of stepsize factor θ :** Our first test is to verify that we can fine-tune the stepsizes by using small pilot run. In this test, we chose between eight different stepsize factors, namely, 0.005, 0.01, 0.05, 0.1, 0.5, 1.0, 5, 10 for both N-SA and E-SA. First, we used short pilot runs ($M = 100$) to select the “most promising” value of the stepsize factor θ , see the beginning of section 5.1. Second, we directly tested which one of the outlined eight values of θ results in the highest quality solution for the sample size $N = 2,000$. The results are presented in the columns “Inferred θ ,” resp., “Best θ ,” of Table 2. As we can see from this table, the inferred θ ’s are very close to the best ones for all test instances and the same conclusion also holds for the E-SA.
- **The effect of changing u :** In Table 3, we report the objective values of EU-1 – EU-5 evaluated at the solutions obtained by N-SA, E-SA and SAA when the sample size is $N = 2,000$. In this table, $\hat{f}(x_*)$ denotes the estimated objective value (using sample size $K = 10,000$) at the obtained solution x_* . Due to the assumption that ξ is normally distributed, the actual objective value $f(x_*)$ can be also computed.

Table 3 Changing u

NAME	N-SA ($\hat{f}(x_*)/f(x_*)$)	E-SA ($\hat{f}(x_*)/f(x_*)$)	SAA ($\hat{f}(x_*)/f(x_*)$)	opt
EU-1	-19.3558/-19.3279	-19.1311/-19.0953	-19.2700/-19.2435	-19.3307
EU-2	-61.4004/-61.3332	-61.7670/-61.6979	-62.8794/-62.7962	-62.9636
EU-3	-81.5215/-81.4339	-80.5735/-80.4873	-83.0845/-82.9732	-83.2145
EU-4	-100.1597/-99.6734	-92.1313/-92.0161	-99.3096/-99.0400	-102.6819
EU-5	-99.5680/-99.2872	-91.2051/-91.0923	-98.5458/-98.2697	-101.9112

Table 4 Changing r

NAME	N-SA($\hat{f}(x_*)/f(x_*)$)	E-SA($\hat{f}(x_*)/f(x_*)$)	SAA($\hat{f}(x_*)/f(x_*)$)	opt
EU-6	-6.2999/-6.2864	-6.2211/-6.2186	-6.3073/-6.3027	-6.3460
EU-7	-16.2514/-16.2294	-15.3818/-15.3717	-16.1474/-16.1226	-16.4738
EU-8	-97.3613/-97.1581	-89.2032/-89.0897	-96.5163/-96.2450	-99.8824
EU-9	-9.540e+2/-9.513e+2	-8.686e+2/-8.675e+2	-9.419e+2/-9.393e+2	-9.757e+2
EU-10	-4.730e+3/-4.717e+3	-4.322e+3/-4.316e+3	-4.689e+3/-4.675e+3	-4.857e+3

Table 5 Lower bounds on optimal values and true optimal values

NAME	N-SA		E-SA		SAA	opt
	f^N	lb^N	f^N	lb^N	\hat{f}_{SAA}^N	
EU-1	-19.4063	-19.2994	-19.4063	-19.2994	-19.4063	-19.3307
EU-2	-62.9984	-62.8754	-62.9984	-62.8758	-62.9984	-62.9367
EU-3	-83.0039	-82.9730	-83.0039	-82.9730	-83.0039	-83.2145
EU-4	-107.5820	-104.5046	-107.2058	-104.4072	-105.0890	-102.6819
EU-5	-107.5745	-104.0644	-108.4063	-104.3577	-104.3214	-101.9112
EU-6	-6.6111	-6.5288	-6.9171	-6.5849	-6.3658	-6.3460
EU-7	-17.0130	-16.7060	-17.1800	-16.7605	-16.7027	-16.4378
EU-8	-106.7958	-102.6311	-106.5921	-102.2588	-102.2914	-99.8824
EU-9	-1029.0530	-997.7217	-1042.7008	-1000.6626	-999.9114	-9.757e2
EU-10	-5192.0409	-4967.9144	-5192.0409	-4981.8515	-4978.2333	-4.857e3

Moreover, a close examination reveals that the optimal value of problem (4.1) can be computed efficiently (see [8]); it is shown in the last column of Table 3.

One interesting observation from this table is that the performance of N-SA is slightly better than that of E-SA even for EU-1 whose feasible region is actually a box instead of a simplex, so that there are no theoretical reasons to prefer N-SA to E-SA.

One other observation from this table is that the solution quality of N-SA significantly outperforms that of E-SA for the two largest values of u . The possible explanation is that the feasible region appears more like a simplex when u is big.

- **The effect of changing r :** Table 4 shows the objective values of EU-6 EU-10 evaluated at the solutions obtained by N-SA, E-SA and SAA when the sample size is $N = 2,000$. In this table, $\hat{f}(x_*)$ and $f(x_*)$, respectively, denote the estimated objective value (using sample size $K = 10,000$) and the actual objective value at the obtained solution x_* , and “opt” denotes the optimal value of problem (4.1).

Recall that the feasible regions for these five instances are simplexes. So, as expected, N-SA consistently outperforms E-SA for all these instances. It is interesting to observe that the objective values achieved by N-SA can be smaller than those by SAA for large r . Note that the SAA problem has been solved to a relatively high accuracy by using the Bundle-level method. For example, for EU-10, the SAA problem was solved to accuracy 0.7e-005.

- **The lower bounds:** Table 5 shows the lower bounds on the objective values of EU-1 – EU-10 obtained by N-SA, E-SA and SAA when the sample size is $N = 2,000$. In Table 5, the lower bounds \underline{f}^N and \underline{lb}^N are the online and offline bounds defined in Section 3. The lower bound for SAA is defined as the optimal value of the corresponding SAA problem. As we can see from this table, the lower bound for SAA is always better than the online lower bound \underline{f}^N for the SA methods (as it should be in the case of constant stepsizes, see Remark 1). However, the offline lower bound \underline{lb}^N can be close or even better than the lower bound obtained from SAA.

Table 6 Variability of the lower bounds for N-SA

NAME	Ind. repl.		Dep. repl.		Whole Traj. \underline{f}^{NM}
	mean	deviation	mean	deviation	
EU-1	-19.5681	0.0857	-19.5387	0.0842	-19.3461
EU-2	-63.3898	0.2372	-63.3786	0.3502	-63.0444
EU-3	-83.6973	0.3121	-83.7339	0.3098	-83.2649
EU-4	-112.2483	1.5616	-114.1652	2.7470	-105.5543
EU-5	-113.7526	1.5951	-115.3103	2.8232	-104.4565
EU-6	-6.7812	0.0265	-6.8969	0.1374	-6.4522
EU-7	-17.7911	0.2326	-18.3881	0.5519	-16.8022
EU-8	-113.5263	2.1348	-117.4176	4.6588	-102.3509
EU-9	-1091.2836	20.2804	-1140.23774	61.1979	-1006.1846
EU-10	-5466.1266	124.5894	-5553.80221	144.6298	-5048.5643

Table 7 Standard deviations

NAME	N-SA		SAA	
	$\hat{f}(x_*)$	$\hat{\sigma}$	$\hat{f}(x_*)$	$\hat{\sigma}$
EU-1	-19.3558	3.1487	-19.2700	3.0019
EU-2	-61.4004	8.4178	-62.8749	8.9099
EU-3	-81.5215	11.7493	-83.0845	12.6015
EU-4	-100.1597	38.6309	-99.3096	61.1053
EU-5	-99.5680	35.1278	-98.5458	60.8440
EU-6	-6.2999	0.6798	-6.3073	0.7030
EU-7	-16.2514	3.5233	-16.1474	5.7941
EU-8	-97.3613	36.3939	-96.5163	61.0974
EU-9	-953.9882	383.8223	-941.9854	611.0414
EU-10	-4729.8534	1746.7144	-4688.9239	3053.7409

Moreover, we estimate the variability of the online lower bounds in the way discussed in section 3.1 and the results are reported in Table 6. In particular, the second and third column of this table show the mean and the standard deviation obtained from $M = 10$ independent replications of N-SA, each of which has the same sample size $N = 1000$. The third and fourth column read the mean and standard deviation computed for the lower bounds associated with the $M = 10$ consecutive partitions of the trajectory of N-SA with a sample size $NM = 10,000$. The last column reports the online lower bound \underline{f}^{NM} . The results indicate that the bounds obtained from independent replications have relatively smaller variability in general.

- **The computation times:** For all instances, the computation times of generating a solution for SA were 10 – 30 times smaller than that for SAA.
- **The standard deviations:** For the generated solution x_* , we evaluate the corresponding objective value $f(x_*)$ by generating an independent large sample ξ_1, \dots, ξ_K , of size $K = 10,000$, and computing the estimate $\hat{f}(x_*) = K^{-1} \sum_{j=1}^K F(x_*, \xi_j)$ of $f(x_*)$. We also computed an estimate of the standard deviation of $F(x_*, \xi)$:

$$\hat{\sigma} = \sqrt{\sum_{j=1}^K (F(x_*, \xi_j) - \hat{f}(x_*))^2 / (K - 1)}.$$

Note that the standard deviation of $\hat{f}(x_*)$, as an estimate of $f(x_*)$, is estimated by $\frac{\hat{\sigma}}{\sqrt{K}}$. Table 7 compares the deviations for N-SA and SAA computed in the above way. From this table, we observe that for instances with either a larger u or larger r , the values of $\hat{\sigma}$ corresponding to the solutions obtained by N-SA can be significantly smaller (up to 1/2) than those by SAA.

5.3 Computational results for the CVaR model

In this subsection, we report some numerical results on applying the Mirror Descent SA method for the CVaR model (4.11). Here the return ξ is assumed to be a normal random vector. In that case random

Table 8 The test instances for CVaR model

NAME	n	β	R	opt
CVaR-1	95	0.05	1.0000	-0.9841
CVaR-2	1,000	0.10	1.0500	1.5272

Table 9 Comparing SA and SAA for the CVaR model

NAME	N	SA					SAA			
		$\hat{f}(x_*)$	$f(x_*)$	\underline{f}_N	\underline{lb}^N	time	$\hat{f}(x_*)$	$f(x_*)$	\hat{f}_{SAA}^N	time
CVaR-1	1000	-0.9807	-0.9823	-1.0695	-1.0136	0	-0.9823	-0.9828	-0.9854	15
	2000	-0.9824	-0.9832	-1.0518	-0.9877	1	-0.9832	-0.9835	-0.9852	27
CVaR-2	1000	1.6048	1.5896	1.1301	1.4590	20	1.6396	1.5795	1.3023	928
	2000	1.5766	1.5633	1.3696	1.4973	39	1.5835	1.5557	1.4780	2784

variable $-\xi^T y$ has normal distribution with mean $-\bar{\xi}^T y$ and variance $y^T \Sigma y$, and

$$\text{CVaR}_{1-\beta}\{-\xi^T y\} = -\bar{\xi}^T y + \rho \sqrt{y^T \Sigma y}, \quad (5.1)$$

where $\rho := \frac{\exp(-z_\beta^2/2)}{\beta\sqrt{2\pi}}$ and $z_\beta := \Phi^{-1}(1 - \beta)$ with $\Phi(\cdot)$ being the cdf of the standard normal distribution. Consequently the optimal solution for (4.11) can be easily obtained by replacing the objective function of (4.11) with the right hand side of (5.1). Clearly, the resulting problem can be reformulated as a conic-quadratic programming program, and its optimal value thus gives us a benchmark to compare the SA and SAA methods.

Two instances for the CVaR model are considered in our experiments. The first instance (CVaR-1) is obtained from [17]. This instance consists of the 95 stocks from S&P100 (excluding SBC, ATI, GS, LU, and VIA-B) and the mean $\bar{\xi}$ and covariance Σ_ξ were estimated using historical monthly prices from 1996 to 2002. The second one (CVaR-2), which contains 1,000 assets, was randomly generated by setting the random return $\xi = \bar{\xi} + Q\zeta$, where ζ is the standard Gaussian vector, $\bar{\xi}_i$ is uniformly distributed in $[0.9, 1.2]$, and Q_{ij} is uniformly distributed in $[0, 0.1]$ for $1 \leq i, j \leq 1,000$. The reliability level β , the bound for expected return R , and the optimal value for these two instances are reported in Table 8.

The computational results for the CVaR model are reported in Table 9, where $\hat{f}(x_*)$ and $f(x_*)$, respectively, denote the estimated objective value (using sample size $K = 10,000$) and the actual objective value at the obtained solution x_* . We conclude from the results in Table 9 that the Mirror Descent SA method can generate good solutions much faster than SAA. The lower bounds derived for the SA method are also comparable to those for the SAA method.

Appendix

We will need the following result (cf., [8, Lemma 6.1]).

Lemma 1 *Let $\zeta_t \in \mathbb{R}^n$, $v_1 \in X^\circ$ and $v_{t+1} = P_{v_t}(\zeta_t)$, $t = 1, \dots, N$. Then*

$$\sum_{t=1}^N \zeta_t^T (v_t - u) \leq V(v_1, u) + (2\alpha)^{-1} \sum_{t=1}^N \|\zeta_t\|_*^2, \quad \forall u \in X. \quad (5.1)$$

We denote here $\delta_t := F(x_t, \xi_t) - f(x_t)$ and $\Delta_t := \mathbf{G}(x_t, \xi_t) - \mathbf{g}(x_t)$. Since x_t is a function of ξ^{t-1} and ξ_t is independent of ξ_t , we have that the conditional expectations

$$\mathbb{E}_{|t-1}[\delta_t] = 0 \quad \text{and} \quad \mathbb{E}_{|t-1}[\Delta_t] = 0, \quad (5.2)$$

and hence the unconditional expectations $\mathbb{E}[\delta_t] = 0$ and $\mathbb{E}[\Delta_t] = 0$ as well.

Part (i) of Theorem 1: Proof. 1.1⁰ If in Lemma 1 we take $v_1 := x_1$ and $\zeta_t := \gamma_t \mathbf{G}(x_t, \xi_t)$, then the corresponding iterates v_t coincide with x_t . Therefore, we have by (5.1) and since $V(x_1, u) \leq D_{\omega, X}^2$ that

$$\sum_{t=1}^N \gamma_t (x_t - u)^T \mathbf{G}(x_t, \xi_t) \leq D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\mathbf{G}(x_t, \xi_t)\|_*^2, \quad \forall u \in X. \quad (5.3)$$

It follows that for any $u \in X$:

$$\begin{aligned} & \sum_{t=1}^N \nu_t [-f(x_t) + (x_t - u)^T \mathbf{g}(x_t)] + \sum_{t=1}^N \nu_t f(x_t) \\ & \leq \frac{D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\mathbf{G}(x_t, \xi_t)\|_*^2}{\sum_{t=1}^N \gamma_t} + \sum_{t=1}^N \nu_t \Delta_t^T (x_t - u). \end{aligned}$$

Since

$$f^{*N} - f_*^N = \sum_{t=1}^N \nu_t f(x_t) + \max_{u \in X} \sum_{t=1}^N \nu_t [-f(x_t) + (x_t - u)^T \mathbf{g}(x_t)],$$

it follows that

$$f^{*N} - f_*^N \leq \frac{D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\mathbf{G}(x_t, \xi_t)\|_*^2}{\sum_{t=1}^N \gamma_t} + \max_{u \in X} \sum_{t=1}^N \nu_t \Delta_t^T (x_t - u). \quad (5.4)$$

Let us estimate the second term in the right hand side of (5.4). Let

$$u_1 = v_1 = x_1; u_{t+1} = P_{u_t}(-\gamma_t \Delta_t), \quad t = 1, 2, \dots, N; v_{t+1} = P_{v_t}(\gamma_t \Delta_t), \quad t = 1, 2, \dots, N. \quad (5.5)$$

Observe that Δ_t is a deterministic function of ξ^t , whence u_t and v_t are deterministic functions of ξ^{t-1} . By using Lemma 1 we obtain

$$\sum_{t=1}^N \gamma_t \Delta_t^T (v_t - u) \leq D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2, \quad \forall u \in X. \quad (5.6)$$

Moreover,

$$\Delta_t^T (v_t - u) = \Delta_t^T (x_t - u) + \Delta_t^T (v_t - x_t),$$

and hence it follows by (5.6) that

$$\max_{u \in X} \sum_{t=1}^N \nu_t \Delta_t^T (x_t - u) \leq \sum_{t=1}^N \nu_t \Delta_t^T (x_t - v_t) + \frac{D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^N \gamma_t}. \quad (5.7)$$

Observe that by similar reasoning applied to $-\Delta_t$ in the role of Δ_t we get

$$\max_{u \in X} \left[-\sum_{t=1}^N \nu_t \Delta_t^T (x_t - u) \right] \leq \left[-\sum_{t=1}^N \nu_t \Delta_t^T (x_t - u_t) \right] + \frac{D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^N \gamma_t}. \quad (5.8)$$

Moreover, $\mathbb{E}_{|t-1}[\Delta_t] = 0$ and u_t, v_t and x_t are functions of ξ^{t-1} , while $\mathbb{E}_{|t-1} \Delta_t = 0$ and hence

$$\mathbb{E}_{|t-1} [(x_t - v_t)^T \Delta_t] = \mathbb{E}_{|t-1} [(x_t - u_t)^T \Delta_t] = 0. \quad (5.9)$$

We also have that $\mathbb{E}_{|t-1} [\|\Delta_t\|_*^2] \leq 4M_*^2$, and hence in view of condition (2.9) it follows from (5.7) and (5.9) that

$$\mathbb{E} \left[\max_{u \in X} \sum_{t=1}^N \nu_t \Delta_t^T (x_t - u) \right] \leq \frac{D_{\omega, X}^2 + 2\alpha^{-1} M_*^2 \sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t}. \quad (5.10)$$

Therefore, by taking expectation of both sides of (5.4) and using (2.9) together with (5.10) we obtain the estimate (3.5).

1.2⁰ In order to prove (3.6) let us observe that $\bar{f}^N - f^{*N} = \sum_{t=1}^N \nu_t \delta_t$, and that for $1 \leq s < t \leq N$,

$$\mathbb{E}[\delta_s \delta_t] = \mathbb{E}\{\mathbb{E}_{|t-1}[\delta_s \delta_t]\} = \mathbb{E}\{\delta_s \mathbb{E}_{|t-1}[\delta_t]\} = 0.$$

Therefore

$$\mathbb{E}\left[(\bar{f}^N - f^{*N})^2\right] = \mathbb{E}\left[\left(\sum_{t=1}^N \nu_t \delta_t\right)^2\right] = \sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\delta_t^2\right] = \sum_{t=1}^N \nu_t^2 \mathbb{E}\left\{\mathbb{E}_{|t-1}\left[\delta_t^2\right]\right\}.$$

Moreover, by condition (2.8) of assumption (A1) we have that $\mathbb{E}_{|t-1}\left[\delta_t^2\right] \leq Q^2$, and hence

$$\mathbb{E}\left[(\bar{f}^N - f^{*N})^2\right] \leq Q^2 \sum_{t=1}^N \nu_t^2. \quad (5.11)$$

Since $\sqrt{\mathbb{E}[Y^2]} \geq \mathbb{E}|Y|$ for any random variable Y , inequality (3.6) follows from (5.11).

1.3⁰ Let us now look at (3.7). We have

$$\begin{aligned} |\underline{f}^N - f_*^N| &= \left| \min_{x \in X} \hat{f}^N(x) - \min_{x \in X} f^N(x) \right| \leq \max_{x \in X} |\hat{f}^N(x) - f^N(x)| \\ &\leq \left| \sum_{t=1}^N \nu_t \delta_t \right| + \max_{x \in X} \left| \sum_{t=1}^N \nu_t \Delta_t^T(x_t - x) \right|. \end{aligned} \quad (5.12)$$

We already showed above (see (5.11)) that

$$\mathbb{E}\left[\left|\sum_{t=1}^N \nu_t \delta_t\right|\right] \leq Q \sqrt{\sum_{t=1}^N \nu_t^2}. \quad (5.13)$$

Invoking (5.7), (5.8), we get

$$\max_{x \in X} \left| \sum_{t=1}^N \nu_t \Delta_t^T(x_t - x) \right| \leq \left| \sum_{t=1}^N \nu_t \Delta_t^T(x_t - v_t) \right| + \left| \sum_{t=1}^N \nu_t \Delta_t^T(x_t - u_t) \right| + \frac{D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^N \gamma_t}. \quad (5.14)$$

Moreover, for $1 \leq s < t \leq N$ we have that $\mathbb{E}\left[(\Delta_s^T(x_s - v_s))(\Delta_t^T(x_t - v_t))\right] = 0$, and hence

$$\begin{aligned} \mathbb{E}\left[\left|\sum_{t=1}^N \nu_t \Delta_t^T(x_t - v_t)\right|^2\right] &= \sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\left|\Delta_t^T(x_t - v_t)\right|^2\right] \leq 4M_*^2 \sum_{t=1}^N \nu_t^2 \mathbb{E}\left[\|x_t - v_t\|^2\right] \\ &\leq 32M_*^2 \alpha^{-1} D_{\omega, X}^2 \sum_{t=1}^N \nu_t^2, \end{aligned}$$

where the last inequality follows by (2.7). It follows that

$$\mathbb{E}\left[\left|\sum_{t=1}^N \nu_t \Delta_t^T(x_t - v_t)\right|\right] \leq 4\sqrt{2\alpha^{-1}} D_{\omega, X} \sqrt{\sum_{t=1}^N \nu_t^2}.$$

By similar reasons,

$$\mathbb{E}\left[\left|\sum_{t=1}^N \nu_t \Delta_t^T(x_t - u_t)\right|\right] \leq 4\sqrt{2\alpha^{-1}} D_{\omega, X} \sqrt{\sum_{t=1}^N \nu_t^2}.$$

These two inequalities combine with (5.13), (5.14) and (5.12) to imply (3.7). This completes the proof of part (i) of Theorem 1. ■

Preparing to prove part (ii) of Theorem 1: To prove part (ii) of Theorem 1 we need the following known result; we give its proof for the sake of completeness.

Lemma 2 Let ξ_1, ξ_2, \dots be a sequence of iid random variables, $\sigma_t > 0$, μ_t , $t = 1, \dots$, be a sequence of deterministic numbers and $\phi_t = \phi_t(\xi^t)$ be deterministic (measurable) functions of $\xi^t = (\xi_1, \dots, \xi_t)$ such that either

Case A: $\mathbb{E}_{|t-1}[\phi_t] = 0$ w.p.1 and $\mathbb{E}_{|t-1}[\exp\{\phi_t^2/\sigma_t^2\}] \leq \exp\{1\}$ w.p.1 for all t , or

Case B: $\mathbb{E}_{|t-1}[\exp\{|\phi_t|/\sigma_t\}] \leq \exp\{1\}$ for all t .

Then for any $\Omega \geq 0$ we have the following. In the case of A:

$$\text{Prob} \left\{ \sum_{t=1}^N \phi_t > \Omega \sqrt{\sum_{t=1}^N \sigma_t^2} \right\} \leq \exp\{-\Omega^2/3\}. \quad (5.15)$$

In the case of B, setting $\sigma^N := (\sigma_1, \dots, \sigma_N)$:

$$\begin{aligned} \text{Prob} \left\{ \sum_{t=1}^N \phi_t > \|\sigma^N\|_1 + \Omega \|\sigma^N\|_2 \right\} &\leq \exp\{-\Omega^2/12\} + \exp\left\{-\frac{3\|\sigma^N\|_2}{4\|\sigma^N\|_\infty} \Omega\right\} \\ &\leq \exp\{-\Omega^2/12\} + \exp\{-3\Omega/4\}. \end{aligned} \quad (5.16)$$

Proof. Let us set $\bar{\phi}_t := \phi_t/\sigma_t$.

Case A: By the respective assumptions about ϕ_t we have that $\mathbb{E}_{|t-1}[\bar{\phi}_t] = 0$ and $\mathbb{E}_{|t-1}[\exp\{\bar{\phi}_t^2\}] \leq \exp\{1\}$ w.p.1. By Jensen inequality it follows that for any $a \in [0, 1]$:

$$\mathbb{E}_{|t-1}[\exp\{a\bar{\phi}_t^2\}] = \mathbb{E}_{|t-1}[(\exp\{\bar{\phi}_t^2\})^a] \leq \left(\mathbb{E}_{|t-1}[\exp\{\bar{\phi}_t^2\}]\right)^a \leq \exp\{a\}.$$

We also have that $\exp\{x\} \leq x + \exp\{9x^2/16\}$ for all x (this can be verified by direct calculations), and hence

$$\mathbb{E}_{|t-1}[\exp\{\lambda\bar{\phi}_t\}] \leq \mathbb{E}_{|t-1}[\exp\{(9\lambda^2/16)\bar{\phi}_t^2\}] \leq \exp\{9\lambda^2/16\}, \quad \forall \lambda \in [0, 4/3]. \quad (5.17)$$

Besides this, we have $\lambda x \leq \frac{3}{8}\lambda^2 + \frac{2}{3}x^2$ for any λ and x , and hence

$$\mathbb{E}_{|t-1}[\exp\{\lambda\bar{\phi}_t\}] \leq \exp\{3\lambda^2/8\} \mathbb{E}_{|t-1}[\exp\{2\bar{\phi}_t^2/3\}] \leq \exp\{2/3 + 3\lambda^2/8\}.$$

Combining the latter inequality with (5.17), we get

$$\mathbb{E}_{|t-1}[\exp\{\lambda\bar{\phi}_t\}] \leq \exp\{3\lambda^2/4\}, \quad \forall \lambda \geq 0.$$

Going back to ϕ_t , the above inequality reads

$$\mathbb{E}_{|t-1}[\exp\{\kappa\phi_t\}] \leq \exp\{3\kappa^2\sigma_t^2/4\}, \quad \forall \kappa \geq 0. \quad (5.18)$$

Now, since ϕ_τ is a deterministic function of ξ^τ and using (5.18), we obtain for any $\kappa \geq 0$:

$$\begin{aligned} \mathbb{E}[\exp\{\kappa \sum_{\tau=1}^t \phi_\tau\}] &= \mathbb{E}[\exp\{\kappa \sum_{\tau=1}^{t-1} \phi_\tau\} \mathbb{E}_{|t-1}[\exp\{\kappa\phi_t\}]] \\ &\leq \exp\{3\kappa^2\sigma_t^2/4\} \mathbb{E}[\exp\{\kappa \sum_{\tau=1}^{t-1} \phi_\tau\}], \end{aligned}$$

and hence

$$\mathbb{E}[\exp\{\kappa \sum_{t=1}^N \phi_t\}] \leq \exp\{3\kappa^2 \sum_{t=1}^N \sigma_t^2/4\}. \quad (5.19)$$

By Markov inequality, we have for $\kappa > 0$ and $\Omega \geq 0$:

$$\begin{aligned} \text{Prob} \left\{ \sum_{t=1}^N \phi_t > \Omega \sqrt{\sum_{t=1}^N \sigma_t^2} \right\} &= \text{Prob} \left\{ \exp \left[\kappa \sum_{t=1}^N \phi_t \right] > \exp \left[\kappa \Omega \sqrt{\sum_{t=1}^N \sigma_t^2} \right] \right\} \\ &\leq \exp \left[-\kappa \Omega \sqrt{\sum_{t=1}^N \sigma_t^2} \right] \mathbb{E} \left\{ \exp \left[\kappa \sum_{t=1}^N \phi_t \right] \right\}. \end{aligned}$$

Together with (5.19) this implies for $\Omega \geq 0$:

$$\text{Prob} \left\{ \sum_{t=1}^N \phi_t > \Omega \sqrt{\sum_{t=1}^N \sigma_t^2} \right\} \leq \inf_{\kappa > 0} \exp \left\{ \frac{3}{4} \kappa^2 \sum_{t=1}^N \sigma_t^2 - \kappa \Omega \sqrt{\sum_{t=1}^N \sigma_t^2} \right\} = \exp \left\{ -\Omega^2/3 \right\}.$$

Case B: Observe first that if η is a random variable such that $\mathbb{E}[\exp\{|\eta|\}] \leq \exp\{1\}$, then

$$0 \leq t \leq \frac{1}{2} \Rightarrow \mathbb{E}[\exp\{t\eta\}] \leq \exp\{t + 3t^2\}. \quad (5.20)$$

Indeed, let $f(t) = \mathbb{E}[\exp\{t\eta\}]$. Then $f(0) = 1$, $f'(0) = \mathbb{E}[\eta] \leq \ln(\mathbb{E}[\exp\{\eta\}]) \leq 1$. Besides this, when $0 \leq t \leq 1/2$, invoking the Cauchy and the Hölder inequalities we have

$$f''(t) = \mathbb{E}[\exp\{t\eta\}\eta^2] \leq [\mathbb{E}[\exp\{2t|\eta|\}]]^{1/2} [\mathbb{E}[\eta^4]]^{1/2} \leq [\mathbb{E}[\exp\{|\eta|\}]]^t [\mathbb{E}[\eta^4]]^{1/2} \leq \exp\{1/2\} [\mathbb{E}[\eta^4]]^{1/2}.$$

It is immediately seen that $s^4 \leq (4/e)^4 \exp\{|s|\}$ for all s , whence $[\mathbb{E}[\eta^4]]^{1/2} \leq (4/e)^2 e^{1/2}$ due to $\mathbb{E}[\exp\{|\eta|\}] \leq e$. Thus, $f''(t) \leq 16/e$ when $0 \leq t \leq 1/2$, and thus $f(t) \leq 1 + t + (8/e)t^2 \leq \exp\{t + (8/e)t^2\} \leq \exp\{t + 3t^2\}$, and (5.20) follows.

Let $\gamma \geq 0$ be such that $\gamma\sigma_t \leq 1/2$, $1 \leq t \leq N$. When $t \leq N$, we have

$$\begin{aligned} \mathbb{E} \left[\exp\left\{ \sum_{\tau=1}^t \gamma \phi_\tau \right\} \right] &= \mathbb{E} \left[\exp\left\{ \sum_{\tau=1}^t \gamma \sigma_\tau \bar{\phi}_\tau \right\} \right] = \mathbb{E} \left[\exp\left\{ \sum_{\tau=1}^{t-1} \gamma \sigma_\tau \bar{\phi}_\tau \right\} \mathbb{E}_{|t-1} \left[\exp\left\{ \gamma \sigma_t \bar{\phi}_t \right\} \right] \right] \\ &\leq \exp\left\{ \gamma \sigma_t + 3\gamma^2 \sigma_t^2 \right\} \mathbb{E} \left[\exp\left\{ \sum_{\tau=1}^{t-1} \gamma \sigma_\tau \bar{\phi}_\tau \right\} \right], \end{aligned}$$

where the concluding inequality is given by (5.20) (note that we are in the case when $\mathbb{E}_{|t-1}[\exp\{|\bar{\phi}_t|\}] \leq \exp\{1\}$ w.p.1). From the resulting recurrence we get

$$0 \leq \gamma \|\sigma^N\|_\infty \leq 1/2 \Rightarrow \mathbb{E} \left[\exp\left\{ \sum_{t=1}^N \gamma \phi_t \right\} \right] \leq \exp\left\{ \gamma \|\sigma^N\|_1 + 3\gamma^2 \|\sigma^N\|_2^2 \right\}.$$

whence for every $\Omega \geq 0$, denoting $\beta_s = \|\sigma^N\|_s$,

$$0 \leq \gamma \beta_\infty \leq 1/2 \Rightarrow p := \text{Prob} \left\{ \sum_{t=1}^N \phi_t > \beta_1 + \Omega \beta_2 \right\} \leq \exp\{3\gamma^2 \beta_2^2 - \gamma \Omega \beta_2\}. \quad (5.21)$$

When $\Omega \leq \bar{\Omega} := 3\beta_2/\beta_\infty$, $\gamma = \Omega/(6\beta_2)$ satisfies the premise in (5.21), and this implication then says that $p \leq \exp\{-\Omega^2/12\}$. When $\Omega > \bar{\Omega}$, we can use the implication with $\gamma = (2\beta_\infty)^{-1}$, thus getting $p \leq \exp\left\{ \frac{\beta_2}{2\beta_\infty} \left[\frac{3\beta_2}{2\beta_\infty} - \Omega \right] \right\} \leq \exp\left\{ -\frac{3\beta_2}{4\beta_\infty} \Omega \right\}$. Thus (5.16) is proved. \blacksquare

Part (ii) of Theorem 1: Proof. Recall that in part (ii) of Theorem 1 assumption (A1) is strengthened to assumption (A2). Then, in addition to (5.2), we have that

$$\mathbb{E}_{|t-1} \left[\exp\{\delta_t^2/Q^2\} \right] \leq \exp\{1\} \quad \text{and} \quad \mathbb{E}_{|t-1} \left[\exp\{\|\Delta_t\|_*^2/(2M_*)^2\} \right] \leq \exp\{1\}. \quad (5.22)$$

Let us also make the following simple observation. If Y_1 and Y_2 are random variables and a_1, a_2, a are numbers such that $a_1 + a_2 \geq a$, then the event $\{Y_1 + Y_2 > a\}$ is included in the union of the events $\{Y_1 > a_1\}$ and $\{Y_2 > a_2\}$, and hence $\text{Prob}\{Y_1 + Y_2 > a\} \leq \text{Prob}\{Y_1 > a_1\} + \text{Prob}\{Y_2 > a_2\}$.

2.1⁰ Recall that $\bar{f}^N - f^{*N} = \sum_{t=1}^N \nu_t \delta_t$, and hence it follows by case A of Lemma 2 together with the first equality in (5.2) and (5.22) that for any $\Omega \geq 0$:

$$\text{Prob} \left\{ \bar{f}^N - f^{*N} > \Omega Q \sqrt{\sum_{t=1}^N \nu_t^2} \right\} \leq \exp\{-\Omega^2/3\}. \quad (5.23)$$

In the same way, by considering $-\delta_t$ instead of δ_t , we have that

$$\text{Prob} \left\{ f^{*N} - \bar{f}^N > \Omega Q \sqrt{\sum_{t=1}^N \nu_t^2} \right\} \leq \exp\{-\Omega^2/3\}, \quad (5.24)$$

The assertion (3.10) follows from (5.23) and (5.24).

2.2⁰ Now by (5.12) and (5.14) we have

$$\left| \underline{f}^N - f_*^N \right| \leq \left| \sum_{t=1}^N \nu_t \delta_t \right| + \left| \sum_{t=1}^N \nu_t \Delta_t^T (x_t - v_t) \right| + \left| \sum_{t=1}^N \nu_t \Delta_t^T (x_t - u_t) \right| + \frac{D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^N \gamma_t}. \quad (5.25)$$

As it was shown above (see (5.23), (5.24)):

$$\text{Prob} \left\{ \left| \sum_{t=1}^N \nu_t \delta_t \right| > \Omega Q \sqrt{\sum_{t=1}^N \nu_t^2} \right\} \leq 2 \exp\{-\Omega^2/3\}. \quad (5.26)$$

Moreover, by (2.7) we have that $\|x_t - v_t\| \leq \|x_t - x_1\| + \|v_t - x_1\| \leq 2\sqrt{2\alpha^{-1}} D_{\omega, X}$, and hence

$$\mathbb{E}_{|t-1} \left[\exp\{|\Delta_t^T (x_t - v_t)|^2 / (4\sqrt{2\alpha^{-1}} D_{\omega, X} M_*)^2\} \right] \leq \exp\{1\}.$$

It follows by case A of Lemma 2 that

$$\text{Prob} \left\{ \left| \sum_{t=1}^N \nu_t \Delta_t^T (x_t - v_t) \right| > 4\Omega \sqrt{2\alpha^{-1}} D_{\omega, X} M_* \sqrt{\sum_{t=1}^N \nu_t^2} \right\} \leq 2 \exp\{-\Omega^2/3\}. \quad (5.27)$$

and similarly

$$\text{Prob} \left\{ \left| \sum_{t=1}^N \nu_t \Delta_t^T (x_t - u_t) \right| > 4\Omega \sqrt{2\alpha^{-1}} D_{\omega, X} M_* \sqrt{\sum_{t=1}^N \nu_t^2} \right\} \leq 2 \exp\{-\Omega^2/3\}. \quad (5.28)$$

Furthermore, invoking (5.22), the random variables $\phi_t = (2\alpha)^{-1} \gamma_t^2 \|\Delta_t\|_*^2 (\sum_{t=1}^N \gamma_t)^{-1}$ satisfy the premise of case B in Lemma 2 with $\sigma_t = 2\alpha^{-1} M_*^2 \gamma_t^2 (\sum_{t=1}^N \gamma_t)^{-1}$. Invoking case B of Lemma, we get

$$\text{Prob} \left\{ \frac{(2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|\Delta_t\|_*^2}{\sum_{t=1}^N \gamma_t} > \frac{2\alpha^{-1} M_*^2 \sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t} + \Omega \frac{2\alpha^{-1} M_*^2 \sqrt{\sum_{t=1}^N \gamma_t^4}}{\sum_{t=1}^N \gamma_t} \right\} \leq \exp\{-\Omega^2/12\} + \exp\{-\Gamma_N \Omega\},$$

$$\Gamma_N = \frac{3\|(\gamma_1^2, \dots, \gamma_N^2)\|_2}{4\|(\gamma_1^2, \dots, \gamma_N^2)\|_\infty} \quad (5.29)$$

Combining this bound with (5.27), (5.28) and taking into account (5.25), we arrive at (3.11).

2.3⁰ It remains to prove (3.9). To this end note by (5.4) and (5.7) we have

$$f^{*N} - f_*^N \leq \frac{2D_{\omega, X}^2 + (2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 (\|G(x_t, \xi_t)\|_*^2 + \|\Delta_t\|_*^2)}{\sum_{t=1}^N \gamma_t} + \sum_{t=1}^N \nu_t \Delta_t^T (x_t - v_t), \quad (5.30)$$

Completely similar to (5.29), we have

$$\text{Prob} \left\{ \frac{(2\alpha)^{-1} \sum_{t=1}^N \gamma_t^2 \|G(x_t, \xi_t)\|_*^2}{\sum_{t=1}^N \gamma_t} > \frac{(2\alpha)^{-1} M_*^2 \sum_{t=1}^N \gamma_t^2}{\sum_{t=1}^N \gamma_t} + \Omega \frac{(2\alpha)^{-1} M_*^2 \sqrt{\sum_{t=1}^N \gamma_t^4}}{\sum_{t=1}^N \gamma_t} \right\} \leq \exp\{-\Omega^2/12\} + \exp\{-\Gamma_N \Omega\} \quad (5.31)$$

This bound combines with (5.29) and (5.27) to imply (3.9). ■

Theorem 2: Proof. Let x_1, \dots, x_N be the trajectory of Mirror Descent SA, and let $x_{N+t} := \tilde{x}_N$, $t = 1, \dots, L$. Then we can write

$$\hat{\ell}_L(x; \tilde{x}_N) = \frac{1}{L} \sum_{t=N+1}^{N+L} \left[F(x_t, \xi_t) + \mathbf{G}(x_t, \xi_t)^T (x - x_t) \right].$$

Let x_* be an optimal solution to (1.1), and let us set $\eta_t := \Delta_t^T (x_* - x_t)$, $t = 1, \dots, N + L$. By (2.7) we have $\|x_t - x_*\| \leq 2\Lambda_{\omega, X}$, and since x_t is a deterministic function of ξ^{t-1} , $1 \leq t \leq N + L$, and the oracle is unbiased, under assumption (A1) we have for $1 \leq t \leq N + L$,

$$\begin{aligned} \mathbb{E}_{|t-1}[\delta_t] &= 0, \mathbb{E}_{|t-1}[\delta_t^2] \leq Q^2, \\ \mathbb{E}_{|t-1}[\eta_t] &= 0, \mathbb{E}_{|t-1}[\eta_t^2] \leq 4\Lambda_{\omega, X}^2 \mathbb{E}_{|t-1}[\|\Delta_t\|_*^2] \leq 16\Lambda_{\omega, X}^2 M_*^2. \end{aligned} \quad (5.32)$$

Consequently

$$\begin{aligned} \hat{f}^N(x_*) &= \underbrace{\frac{1}{N} \sum_{t=1}^N [f(x_t) + \mathbf{g}(x_t)^T (x_* - x_t)]}_{\leq f(x_*) = \text{Opt}} + \underbrace{\frac{1}{N} \sum_{t=1}^N [\delta_t + \eta_t]}_{\zeta_1}, \\ \hat{\ell}_L(x_*; \tilde{x}_N) &= \underbrace{\frac{1}{L} \sum_{t=N+1}^{N+L} [f(x_t) + \mathbf{g}(x_t)^T (x_* - x_t)]}_{\leq f(x_*) = \text{Opt}} + \underbrace{\frac{1}{L} \sum_{t=N+1}^{N+L} [\delta_t + \eta_t]}_{\zeta_2}. \end{aligned}$$

It follows that

$$lb^N - \text{Opt} \leq \max\{\hat{f}^N(x_*), \hat{\ell}_L(x_*; \tilde{x}_N)\} - \text{Opt} \leq \max\{\zeta_1, \zeta_2\} \leq |\zeta_1| + |\zeta_2|. \quad (5.33)$$

From (5.32) it follows that

$$\begin{aligned} \mathbb{E}[\zeta_1^2] &\leq N^{-1} \left(2\mathbb{E}[\delta_t^2] + 2\mathbb{E}[\eta_t^2] \right) \leq (2Q^2 + 32\Lambda_{\omega, X}^2 M_*^2) N^{-1}, \\ \mathbb{E}[\zeta_2^2] &\leq L^{-1} \left(2\mathbb{E}[\delta_t^2] + 2\mathbb{E}[\eta_t^2] \right) \leq (2Q^2 + 32\Lambda_{\omega, X}^2 M_*^2) L^{-1}, \end{aligned}$$

which combines with (5.33) to imply (3.15).

Under assumption (A2), along with (5.32) we also have that

$$\mathbb{E}_{|t-1} \left[\exp\{\delta_t^2/Q^2\} \right] \leq \exp\{1\}, \quad \mathbb{E}_{|t-1} \left[\exp\{\eta_t^2/(4\Lambda_{\omega, X} M_*)^2\} \right] \leq \exp\{1\},$$

and hence

$$\mathbb{E}_{|t-1} [\delta_t + \eta_t] = 0, \quad \mathbb{E}_{|t-1} \left[\exp\{[\delta_t + \eta_t]^2/(Q + 4\Lambda_{\omega, X} M_*)^2\} \right] \leq \exp\{1\}.$$

Invoking case A of Lemma 2, we conclude that for all $\Omega \geq 0$:

$$\begin{aligned} \text{Prob} \left\{ |\zeta_1| > \Omega[Q + 4\Lambda_{\omega, X} M_*] N^{-1/2} \right\} &\leq 2 \exp\{-\Omega^2/3\}, \\ \text{Prob} \left\{ |\zeta_2| > [Q + 4\Lambda_{\omega, X} M_*] L^{-1/2} \right\} &\leq 2 \exp\{-\Omega^2/3\}, \end{aligned}$$

which combines with (5.33) to imply (3.16). ■

References

1. Bregman, L.M., The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming, *Comput. Math. Math. Phys.*, **7**, 200-217 (1967).
2. Kleywegt, A. J., Shapiro, A. and Homem-de-Mello, T., The sample average approximation method for stochastic discrete optimization, *SIAM J. Optimization*, **12**, 479-502 (2001).
3. Lemarechal, C., Nemirovski, A., and Nesterov, Yu., New variants of bundle methods, *Math. Progr.* **69**, 111-148 (1995).
4. Linderoth, J., Shapiro, A. and Wright, S., The empirical behavior of sampling methods for stochastic programming, *Annals of Oper. Res.*, **142**, 215-241 (2006).
5. Mak, W.K., Morton, D.P. and Wood, R.K., Monte Carlo bounding techniques for determining solution quality in stochastic programs, *Oper. Res. Letters*, **24**, 47-56 (1999).
6. Nemirovskii, A., and Yudin, D. "On Cezari's convergence of the steepest descent method for approximating saddle point of convex-concave functions." (in Russian) - *Doklady Akademii Nauk SSSR*, **239:5** (1978) (English translation: *Soviet Math. Dokl.* **19:2** (1978))
7. Nemirovski, A., Yudin, D., *Problem complexity and method efficiency in optimization*, Wiley-Interscience Series in Discrete Mathematics, John Wiley, XV, 1983.
8. Nemirovski, A., Juditsky, A., Lan, G., and Shapiro, A., Robust stochastic approximation approach to stochastic programming, *SIAM J. Optimization*, **19**, 1574-1609 (2009).
9. Norikin, V. I., Pflug, G. Ch. and A. Ruszczyński, A., A branch and bound method for stochastic global optimization, *Math. Progr.*, **83**, 425-450 (1998).
10. Polyak, B.T., New stochastic approximation type procedures, *Automat. i Telemekh.* **7** (1990), 98-107 (English translation: *Automation and Remote Control*).
11. Polyak, B.T. and Juditsky, A.B., Acceleration of stochastic approximation by averaging, *SIAM J. Contr. and Optim.*, **30**, 838-855 (1992).
12. Robbins, H. and Monro, S., A stochastic approximation method, *Annals of Math. Stat.*, **22**, 400-407 (1951).
13. Rockafellar, R.T. and Uryasev, S.P., Optimization of conditional value-at-risk, *The Journal of Risk*, **2**, 21-41 (2000).
14. Shapiro, A., Monte Carlo sampling methods, in: Ruszczyński, A. and Shapiro, A., (Eds.), *Stochastic Programming*, Handbook in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003.
15. Shapiro, A. and Nemirovski, A., On complexity of stochastic programming problems, in: V. Jeyakumar and A.M. Rubinov (Eds.), *Continuous Optimization: Current Trends and Applications*, 111-144, Springer, 2005.
16. Verweij, B., Ahmed, S., Kleywegt, A.J., Nemhauser, G. and Shapiro, A., The sample average approximation method applied to stochastic routing problems: a computational study, *Computational Optimization and Applications*, **24**, 289-333 (2003).
17. Wang W. and Ahmed, S., Sample average approximation of expected value constrained stochastic programs, *E-print available at: <http://www.optimization-online.org>*, 2007.

Table 10 SA vs SAA for EU-1

		VERIFICATION				CONSTRUCTION		
ALG.	step	obj	dev	\bar{lb}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-19.3503	3.1475	-18.9132	-18.9132	-20.6156	-19.2006	0
	0.010	-19.3420	3.1445	-18.9131	-18.9131	-20.6156	-19.1929	1
	0.050	-19.2762	3.1205	-18.9127	-18.9127	-20.6156	-19.1315	0
	0.100	-19.1952	3.0915	-18.9124	-18.9124	-20.6156	-19.0559	0
	0.500	-18.5939	2.8906	-18.9112	-18.9112	-20.6156	-18.4886	1
	1.000	-17.9408	2.6962	-18.9115	-18.9115	-20.6156	-17.8641	0
	5.000	-14.8128	1.9667	-18.9265	-18.9265	-20.6156	-14.8226	0
	10.000	-13.0173	1.6291	-18.9382	-18.9382	-20.6156	-13.0614	1
N-SA (N=1000)	0.005	-19.3549	3.1484	-19.2891	-19.2863	-19.4403	-19.1874	2
	0.010	-19.3512	3.1463	-19.2891	-19.2863	-19.4403	-19.1838	2
	0.050	-19.3219	3.1293	-19.2891	-19.2862	-19.4403	-19.1552	2
	0.100	-19.2856	3.1088	-19.2891	-19.2862	-19.4403	-19.1197	2
	0.500	-19.0087	2.9659	-19.2891	-19.2861	-19.4403	-18.8495	2
	1.000	-18.6918	2.8258	-19.2892	-19.2862	-19.4403	-18.5398	2
	5.000	-16.8467	2.2564	-19.2909	-19.2876	-19.4403	-16.7191	2
	10.000	-15.3926	1.9396	-19.2923	-19.2888	-19.4403	-15.2802	2
N-SA (N=2000)	0.005	-19.3558	3.1487	-19.2998	-19.2994	-19.4063	-19.2671	3
	0.010	-19.3530	3.1468	-19.2998	-19.2994	-19.4063	-19.2643	3
	0.050	-19.3310	3.1319	-19.2998	-19.2994	-19.4063	-19.2426	3
	0.100	-19.3038	3.1139	-19.2998	-19.2994	-19.4063	-19.2156	3
	0.500	-19.0948	2.9879	-19.2998	-19.2994	-19.4063	-19.0088	3
	1.000	-18.8530	2.8640	-19.2999	-19.2995	-19.4063	-18.7693	4
	5.000	-17.4311	2.3983	-19.3011	-19.3007	-19.4063	-17.3062	3
	10.000	-16.2469	2.1547	-19.3021	-19.3016	-19.4063	-16.0626	3
E-SA (N=100)	0.005	-18.6749	2.9086	-18.9113	-18.9113	-20.6156	-18.5716	0
	0.010	-17.9962	2.6935	-18.9123	-18.9123	-20.6156	-17.9301	0
	0.050	-15.1313	2.0948	-18.9150	-18.9150	-20.6156	-15.1044	0
	0.100	-14.1627	1.9553	-18.9134	-18.9134	-20.6156	-14.1330	0
	0.500	-13.2785	1.8369	-18.9124	-18.9124	-20.6156	-13.3260	1
	1.000	-13.1712	1.8221	-18.9122	-18.9122	-20.6156	-13.2404	0
	5.000	-13.0722	1.8108	-18.9122	-18.9122	-20.6156	-13.1487	1
	10.000	-13.0589	1.8091	-18.9122	-18.9122	-20.6156	-13.1320	0
E-SA (N=1000)	0.005	-19.0549	2.9808	-19.2891	-19.2861	-19.4403	-18.8944	1
	0.010	-18.7532	2.8325	-19.2893	-19.2862	-19.4403	-18.6035	1
	0.050	-16.9969	2.3496	-19.2898	-19.2867	-19.4403	-16.8469	1
	0.100	-15.7783	2.3432	-19.2896	-19.2866	-19.4403	-15.6116	2
	0.500	-13.6786	1.8912	-19.2893	-19.2864	-19.4403	-13.5487	1
	1.000	-13.3306	1.8349	-19.2892	-19.2863	-19.4403	-13.2329	1
	5.000	-13.0408	1.8000	-19.2892	-19.2863	-19.4403	-12.9527	1
	10.000	-13.0024	1.7957	-19.2892	-19.2863	-19.4403	-12.9122	1
E-SA (N=2000)	0.005	-19.1311	3.0016	-19.2998	-19.2994	-19.4063	-19.0447	2
	0.010	-18.9050	2.8701	-19.3000	-19.2995	-19.4063	-18.8221	1
	0.050	-17.5945	2.5185	-19.3003	-19.2999	-19.4063	-17.4574	2
	0.100	-16.5355	2.3273	-19.3003	-19.2999	-19.4063	-16.3353	2
	0.500	-14.4868	2.3222	-19.3000	-19.2996	-19.4063	-13.8779	2
	1.000	-14.1098	2.3486	-19.2999	-19.2996	-19.4063	-13.4123	2
	5.000	-13.5949	2.3780	-19.2999	-19.2995	-19.4063	-13.0214	2
	10.000	-13.4814	2.3808	-19.2999	-19.2995	-19.4063	-12.9633	2
SAA	N=100	-18.5799	2.8127	-	-	-20.6156	-20.6156	2
SAA	N=1000	-19.2104	2.9673	-	-	-19.4403	-19.4403	14
SAA	N=2000	-19.2700	3.0019	-	-	-19.4063	-19.4063	27

Table 11 SA vs SAA for EU-2

		VERIFICATION				CONSTRUCTION		
ALG.	step	obj	dev	\bar{b}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-38.1358	6.3013	-62.0678	-62.0679	-67.3351	-37.7797	1
	0.010	-38.2777	6.3016	-62.0637	-62.0637	-67.3351	-37.8657	1
	0.050	-39.5633	6.3170	-62.0320	-62.0320	-67.3351	-38.5604	1
	0.100	-41.3122	6.3023	-61.9965	-61.9965	-67.3351	-39.4440	1
	0.500	-51.5592	7.1720	-61.8789	-61.8789	-67.3351	-45.6782	1
	1.000	-53.7202	7.5889	-61.8677	-61.8677	-67.3351	-48.6930	0
	5.000	-50.1970	7.2056	-61.8935	-61.8935	-67.3351	-48.9589	0
	10.000	-44.3844	6.2152	-61.8906	-61.8906	-67.3351	-44.6652	0
N-SA (N=1000)	0.005	-38.5134	6.3022	-62.8773	-62.8591	-63.0584	-37.9218	3
	0.010	-39.0381	6.3052	-62.8761	-62.8583	-63.0584	-38.1864	3
	0.050	-43.4184	6.4096	-62.8675	-62.8524	-63.0584	-40.3681	3
	0.100	-49.1950	6.7643	-62.8607	-62.8477	-63.0584	-43.2200	3
	0.500	-59.5124	7.9965	-62.8492	-62.8396	-63.0584	-53.4417	3
	1.000	-60.5022	8.2671	-62.8475	-62.8383	-63.0584	-56.0828	3
	5.000	-57.4415	8.0493	-62.8512	-62.8410	-63.0584	-55.6279	3
	10.000	-54.4349	7.6609	-62.8532	-62.8425	-63.0584	-53.1233	3
N-SA (N=2000)	0.005	-38.7364	6.3032	-62.8993	-62.8945	-62.9984	-38.1877	6
	0.010	-39.4895	6.3093	-62.8981	-62.8935	-62.9984	-38.5625	5
	0.050	-45.8521	6.5226	-62.8898	-62.8866	-62.9984	-41.6791	5
	0.100	-52.7069	7.0492	-62.8850	-62.8825	-62.9984	-45.5742	6
	0.500	-60.8674	8.2380	-62.8776	-62.8762	-62.9984	-55.6583	5
	1.000	-61.4004	8.4178	-62.8768	-62.8754	-62.9984	-57.9291	6
	5.000	-58.9613	8.1694	-62.8790	-62.8775	-62.9984	-57.5192	5
	10.000	-56.5552	8.1115	-62.8805	-62.8788	-62.9984	-55.3438	5
E-SA (N=100)	0.005	-44.7298	6.5411	-61.9388	-61.9387	-67.3351	-41.2221	0
	0.010	-50.8755	7.1191	-61.8813	-61.8813	-67.3351	-44.6372	0
	0.050	-52.9844	7.7711	-61.8799	-61.8799	-67.3351	-50.4057	1
	0.100	-48.9209	7.0075	-61.9036	-61.9036	-67.3351	-48.3108	1
	0.500	-43.1763	6.4523	-61.9653	-61.9653	-67.3351	-43.2668	0
	1.000	-42.4103	6.4126	-61.9776	-61.9776	-67.3351	-42.6107	0
	5.000	-41.7427	6.3835	-61.9889	-61.9889	-67.3351	-42.1865	1
	10.000	-41.6408	6.3801	-61.9907	-61.9907	-67.3351	-42.0068	1
E-SA (N=1000)	0.005	-56.4912	7.5038	-62.8524	-62.8419	-63.0584	-47.9520	2
	0.010	-60.6441	8.1788	-62.8485	-62.8390	-63.0584	-53.2857	1
	0.050	-59.4372	8.1330	-62.8483	-62.8389	-63.0584	-57.1628	2
	0.100	-56.4664	7.8226	-62.8509	-62.8408	-63.0584	-55.1036	1
	0.500	-47.1630	6.9548	-62.8617	-62.8485	-63.0584	-46.1799	2
	1.000	-44.2730	6.6700	-62.8657	-62.8512	-63.0584	-43.7097	1
	5.000	-42.1794	6.4795	-62.8697	-62.8539	-63.0584	-41.8224	2
	10.000	-41.9037	6.4625	-62.8702	-62.8543	-63.0584	-41.5539	1
E-SA (N=2000)	0.005	-59.5578	7.9151	-62.8792	-62.8776	-62.9984	-51.0444	3
	0.010	-61.7670	8.4029	-62.8772	-62.8758	-62.9984	-55.8117	4
	0.050	-60.3580	8.2543	-62.8773	-62.8759	-62.9984	-58.7962	3
	0.100	-58.0879	7.9671	-62.8788	-62.8773	-62.9984	-57.1121	3
	0.500	-48.9764	7.1088	-62.8862	-62.8836	-62.9984	-48.2204	3
	1.000	-45.8957	7.8486	-62.8898	-62.8865	-62.9984	-45.0630	4
	5.000	-43.3615	8.5239	-62.8937	-62.8898	-62.9984	-42.0979	3
	10.000	-43.1312	8.7238	-62.8943	-62.8903	-62.9984	-41.7328	3
SAA	N=100	-58.9225	8.9102	-	-	-67.3351	-67.3350	2
SAA	N=1000	-62.6459	8.9095	-	-	-63.0584	-63.0584	19
SAA	N=2000	-62.8749	8.9099	-	-	-62.9984	-62.9983	38

Table 12 SA vs SAA for EU-3

ALG.	step	VERIFICATION				CONSTRUCTION		
		obj	dev	\bar{lb}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-39.6117	6.3013	-82.3157	-82.3157	-88.4304	-39.2556	1
	0.010	-39.7537	6.3016	-82.3084	-82.3084	-88.4304	-39.3416	1
	0.050	-41.0393	6.3170	-82.2505	-82.2505	-88.4304	-40.0363	0
	0.100	-42.7881	6.3623	-82.1798	-82.1798	-88.4304	-40.9200	0
	0.500	-57.4667	7.8849	-81.8238	-81.8239	-88.4304	-48.3655	0
	1.000	-67.3666	9.5826	-81.7461	-81.7461	-88.4304	-55.4946	0
	5.000	-68.6110	10.8605	-81.7913	-81.7913	-88.4304	-63.1347	1
	10.000	-65.6632	11.2328	-81.8150	-81.8150	-88.4304	-62.5604	1
N-SA (N=1000)	0.005	-39.9893	6.3022	-83.1274	-82.9784	-83.1184	-39.3977	3
	0.010	-40.5141	6.3052	-83.1252	-82.9773	-83.1184	-39.6624	2
	0.050	-44.8943	6.4096	-83.1080	-82.9691	-83.1184	-41.8440	3
	0.100	-50.7953	6.7777	-83.0894	-82.9602	-83.1184	-44.7043	3
	0.500	-77.0042	10.6584	-83.0562	-82.9434	-83.1184	-63.0870	3
	1.000	-80.2141	11.4031	-83.0530	-82.9415	-83.1184	-70.4144	2
	5.000	-77.8238	11.1152	-83.0589	-82.9448	-83.1184	-74.4543	3
	10.000	-75.5959	11.1703	-83.0627	-82.9468	-83.1184	-72.6843	3
N-SA (N=2000)	0.005	-40.2123	6.3032	-83.1493	-82.9893	-83.0039	-39.6637	5
	0.010	-40.9655	6.3093	-83.1470	-82.9888	-83.0039	-40.0384	5
	0.050	-47.3281	6.5226	-83.1294	-82.9845	-83.0039	-43.1550	6
	0.100	-55.7484	7.2340	-83.1132	-82.9798	-83.0039	-47.2583	5
	0.500	-80.1848	11.3168	-83.0944	-82.9737	-83.0039	-67.8061	6
	1.000	-81.5215	11.7493	-83.0932	-82.9730	-83.0039	-73.7930	5
	5.000	-79.2787	11.3618	-83.0972	-82.9747	-83.0039	-76.7361	5
	10.000	-77.1648	10.9266	-83.0998	-82.9755	-83.0039	-75.1356	5
E-SA (N=100)	0.005	-46.2058	6.5411	-82.0663	-82.0663	-88.4304	-42.6980	0
	0.010	-52.9706	7.1919	-81.9096	-81.9096	-88.4304	-46.2106	0
	0.050	-69.0513	10.0283	-81.7446	-81.7446	-88.4304	-59.6579	1
	0.100	-64.4670	9.7883	-81.8014	-81.8014	-88.4304	-60.0157	1
	0.500	-51.4511	10.2614	-81.9668	-81.9668	-88.4304	-49.8357	0
	1.000	-49.3325	11.0861	-82.0287	-82.0287	-88.4304	-46.8866	0
	5.000	-45.5870	6.5358	-82.0956	-82.0956	-88.4304	-43.3252	1
	10.000	-45.3499	6.5216	-82.1024	-82.1024	-88.4304	-43.1719	1
E-SA (N=1000)	0.005	-60.9010	7.8322	-83.0739	-82.9526	-83.1184	-49.9017	2
	0.010	-74.4544	9.9248	-83.0616	-82.9463	-83.1184	-59.1547	2
	0.050	-79.1243	11.3623	-83.0530	-82.9414	-83.1184	-73.2204	2
	0.100	-75.8435	10.8148	-83.0551	-82.9428	-83.1184	-72.6632	2
	0.500	-60.2684	9.7266	-83.0720	-82.9515	-83.1184	-58.7534	1
	1.000	-53.4428	9.7788	-83.0846	-82.9579	-83.1184	-52.3161	2
	5.000	-46.7625	6.4998	-83.1021	-82.9663	-83.1184	-46.4668	2
	10.000	-45.9571	6.4594	-83.1048	-83.1048	-83.9029	-45.5528	2
E-SA (N=2000)	0.005	-68.6529	8.8532	-83.1047	-82.9771	-83.0039	-54.3649	3
	0.010	-78.6518	10.7185	-83.0975	-82.9748	-83.0039	-64.4228	3
	0.050	-80.5735	11.6903	-83.0931	-82.9730	-83.0039	-75.9516	3
	0.100	-78.2839	11.2592	-83.0945	-82.9736	-83.0039	-75.4904	3
	0.500	-64.3056	9.2455	-83.1041	-82.9769	-83.0039	-62.9399	3
	1.000	-56.8833	9.8019	-83.1131	-82.9798	-83.0039	-55.4457	4
	5.000	-48.9097	11.1414	-83.1294	-82.9845	-83.0039	-47.0649	3
	10.000	-47.8454	11.8443	-83.1323	-83.1322	-83.3646	-45.8304	3
SAA	N=100	-77.8239	12.5917	-	-	-88.4304	-88.4304	3
SAA	N=1000	-82.8458	12.6014	-	-	-83.1184	-83.1183	20
SAA	N=2000	-83.0845	12.6015	-	-	-83.0039	-83.0039	38

Table 13 SA vs SAA for EU-4

		VERIFICATION				CONSTRUCTION		
ALG.	step	obj	dev	\bar{lb}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-39.1382	6.3012	-103.8668	-103.8668	-135.1012	-38.8100	1
	0.010	-39.2088	6.3013	-103.8612	-103.8612	-135.1012	-38.8528	0
	0.050	-39.7787	6.3059	-103.8158	-103.8156	-135.1012	-39.1969	0
	0.100	-40.6311	6.3169	-103.7576	-103.7575	-135.1012	-39.6309	0
	0.500	-47.8499	6.6972	-103.2766	-103.2703	-135.1012	-43.2413	0
	1.000	-57.3197	8.0030	-102.8337	-102.7924	-135.1012	-47.9764	1
	5.000	-86.1072	51.7058	-103.4591	-103.4589	-156.8338	-70.8437	1
	10.000	-90.2683	34.5832	-103.7156	-103.7169	-156.8338	-82.7600	1
N-SA (N=1000)	0.005	-39.3257	6.3014	-104.6774	-104.5178	-110.8567	-38.8632	3
	0.010	-39.5851	6.3022	-104.6757	-104.5164	-110.8567	-38.9944	3
	0.050	-41.7089	6.3270	-104.6621	-104.5034	-110.8567	-40.0605	3
	0.100	-44.4731	6.4088	-104.6440	-104.4866	-110.8567	-41.4323	3
	0.500	-67.7814	9.2616	-104.5130	-104.3583	-110.8567	-53.1137	3
	1.000	-86.1106	15.0456	-104.4343	-104.2763	-110.8567	-65.5627	3
	5.000	-99.5512	38.4448	-104.4886	-104.4763	-112.8418	-89.6211	3
	10.000	-98.4015	41.0811	-104.4814	-104.4814	-126.7905	-90.2347	3
N-SA (N=2000)	0.005	-39.4361	6.3017	-104.6990	-104.5673	-107.2058	-39.0746	5
	0.010	-39.8073	6.3032	-104.6976	-104.5658	-107.2058	-39.2600	5
	0.050	-42.8712	6.3533	-104.6831	-104.5530	-107.2058	-40.7748	5
	0.100	-46.8982	6.5210	-104.6635	-104.5357	-107.2058	-42.7393	5
	0.500	-76.9801	11.2353	-104.5415	-104.4233	-107.2058	-58.7447	6
	1.000	-92.4109	18.0166	-104.4804	-104.3618	-107.2058	-72.2851	6
	5.000	-100.1597	38.6309	-104.6123	-104.5046	-107.5820	-92.0804	6
	10.000	-98.7639	33.4377	-104.7341	-104.7341	-115.3887	-92.4497	5
E-SA (N=100)	0.005	-42.3324	6.3614	-103.6555	-103.6555	-135.1012	-40.5109	0
	0.010	-45.7234	6.5356	-103.4658	-103.4658	-135.1012	-42.2545	0
	0.050	-65.4026	9.1650	-102.8399	-102.8251	-135.1012	-54.0087	0
	0.100	-71.9510	11.0358	-102.5784	-102.5595	-135.1012	-60.2342	0
	0.500	-64.6265	11.8169	-103.0449	-102.6408	-135.1012	-61.9645	1
	1.000	-58.5949	13.5084	-102.9164	-102.5881	-135.1012	-57.2348	1
	5.000	-51.8700	7.5180	-103.0982	-103.0799	-135.1012	-51.9177	1
	10.000	-51.4719	8.1276	-103.3523	-103.3523	-244.1578	-46.3071	1
E-SA (N=1000)	0.005	-49.6303	6.7088	-104.6225	-104.4666	-110.8567	-44.0540	2
	0.010	-60.2552	7.8013	-104.5894	-104.4354	-110.8567	-49.3742	2
	0.050	-84.8213	12.8601	-104.5060	-104.3555	-110.8567	-71.1002	2
	0.100	-89.1139	15.4349	-104.4634	-104.3129	-110.8567	-78.0858	2
	0.500	-84.9472	18.8111	-104.3991	-104.2318	-110.8567	-80.3334	1
	1.000	-75.0741	14.9765	-104.4396	-104.2645	-110.8567	-73.2758	2
	5.000	-55.2427	7.3760	-104.5551	-104.4008	-110.8567	-54.8496	2
	10.000	-52.8571	7.0963	-104.5760	-104.5760	-126.2563	-51.0361	2
E-SA (N=2000)	0.005	-54.0474	7.0659	-104.6468	-104.5213	-107.2058	-46.3947	3
	0.010	-68.0061	8.8163	-104.6229	-104.5002	-107.2058	-53.7987	3
	0.050	-88.4424	14.1327	-104.5570	-104.4399	-107.2058	-75.8246	3
	0.100	-92.1313	16.8344	-104.5233	-104.4072	-107.2058	-82.1191	4
	0.500	-89.9439	21.9351	-104.4546	-104.3297	-107.2058	-84.7472	3
	1.000	-81.7297	19.6360	-104.4727	-104.3299	-106.7333	-77.5126	3
	5.000	-57.7183	8.8114	-104.5973	-104.4756	-107.2058	-56.7866	3
	10.000	-55.4649	15.4939	-104.6239	-104.5875	-109.0572	-52.8853	3
SAA	N=100	-72.3744	143.0249	-	-	-134.5648	-134.5647	3
SAA	N=1000	-96.2697	72.8944	-	-	-108.3190	-108.3142	93
SAA	N=2000	-99.3096	61.1053	-	-	-105.0890	-105.0881	163

Table 14 SA vs SAA for EU-5

ALG.	step	VERIFICATION				CONSTRUCTION		
		obj	dev	\bar{lb}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-38.3300	6.3012	-103.8925	-103.8925	-146.7304	-38.0152	1
	0.010	-38.3669	6.3012	-103.8890	-103.8890	-146.7304	-38.0375	1
	0.050	-38.6626	6.3019	-103.8610	-103.8610	-146.7304	-38.2167	1
	0.100	-39.0423	6.3062	-103.8250	-103.8250	-146.7304	-38.4418	0
	0.500	-42.6895	6.4052	-103.5194	-103.5183	-146.7304	-40.2841	0
	1.000	-47.4963	6.7351	-103.2526	-103.2406	-146.7304	-42.6755	0
	5.000	-79.9383	18.7838	-103.1278	-102.9631	-146.7304	-61.4136	0
	10.000	-82.0792	34.7555	-104.2617	-104.2619	-279.7724	-52.0580	0
N-SA (N=1000)	0.005	-38.4278	6.3012	-104.7014	-104.5825	-112.9030	-38.0263	3
	0.010	-38.5629	6.3014	-104.7006	-104.5813	-112.9030	-38.0947	3
	0.050	-39.6565	6.3081	-104.6938	-104.5749	-112.9030	-38.6461	3
	0.100	-41.0560	6.3294	-104.6840	-104.5659	-112.9030	-39.3467	3
	0.500	-53.2736	7.1266	-104.6034	-104.4865	-112.9030	-45.3025	3
	1.000	-68.1682	9.5052	-104.5172	-104.3987	-112.9030	-52.9831	3
	5.000	-97.9097	33.0506	-104.5134	-104.1638	-112.9030	-82.6433	3
	10.000	-97.3159	38.5009	-104.4859	-104.4859	-123.3988	-85.8145	3
N-SA (N=2000)	0.005	-38.4854	6.3013	-104.7198	-104.5355	-108.4063	-38.2120	5
	0.010	-38.6783	6.3017	-104.7228	-104.5320	-108.4063	-38.3085	5
	0.050	-40.2483	6.3150	-104.7155	-104.5281	-108.4063	-39.0894	5
	0.100	-42.2733	6.3583	-104.7058	-104.5187	-108.4063	-40.0870	5
	0.500	-59.6215	7.8891	-104.6267	-104.4384	-108.4063	-48.5978	6
	1.000	-77.4046	11.5699	-104.5631	-104.3676	-108.4063	-58.7659	5
	5.000	-99.5680	35.1278	-104.5241	-104.0644	-107.5745	-86.5396	6
	10.000	-98.9885	34.4152	-104.9583	-104.9291	-109.5558	-90.8342	5
E-SA (N=100)	0.005	-41.4236	6.3568	-103.6469	-103.6468	-146.7304	-39.6674	1
	0.010	-44.6804	6.5178	-103.4361	-103.4354	-146.7304	-41.3419	1
	0.050	-64.1034	9.0602	-102.8722	-102.8181	-146.7304	-52.8315	0
	0.100	-70.9505	10.9285	-102.7239	-102.5757	-146.7304	-59.1516	0
	0.500	-64.2812	11.8766	-103.3224	-102.7288	-146.7304	-61.4639	1
	1.000	-58.1064	13.4753	-103.5058	-102.7123	-146.7304	-56.6712	1
	5.000	-51.1580	7.5291	-103.4436	-103.4436	-146.7304	-51.2048	0
	10.000	-51.2166	13.0679	-103.8111	-103.8111	-261.9004	-45.2895	0
E-SA (N=1000)	0.005	-48.4335	6.6780	-104.6471	-104.5295	-112.9030	-43.0687	2
	0.010	-58.6551	7.6979	-104.6129	-104.4961	-112.9030	-48.1791	2
	0.050	-83.6780	12.7125	-104.5254	-104.4092	-112.9030	-69.8393	1
	0.100	-88.1730	15.2741	-104.4785	-104.3620	-112.9030	-76.9825	2
	0.500	-84.5810	18.8582	-104.4300	-104.2320	-112.9030	-79.8171	1
	1.000	-74.9702	15.3083	-104.4912	-104.2363	-112.9030	-73.0591	2
	5.000	-54.6844	7.4084	-104.5690	-104.4563	-112.9030	-54.2912	1
	10.000	-52.2596	7.1116	-104.4963	-104.4963	-127.0456	-50.5816	2
E-SA (N=2000)	0.005	-52.6794	7.0095	-104.6720	-104.4858	-108.4063	-45.3230	3
	0.010	-66.3654	8.6960	-104.6477	-104.4615	-108.4063	-52.4664	3
	0.050	-87.3763	13.9872	-104.5820	-104.3943	-108.4063	-74.6040	3
	0.100	-91.2051	16.6639	-104.5481	-104.3577	-108.4063	-81.0510	4
	0.500	-89.4774	21.9242	-104.4849	-104.2575	-108.4063	-84.1870	3
	1.000	-81.5608	19.8583	-104.4922	-104.2144	-107.5745	-77.2696	3
	5.000	-57.3049	8.9504	-104.5317	-104.4206	-108.4811	-56.3306	3
	10.000	-54.7480	15.3229	-104.6310	-104.5285	-111.2797	-52.2968	4
SAA	N=100	16.3822	352.8689	-	-	-139.9952	-139.9946	7
SAA	N=1000	-95.5139	72.8589	-	-	-107.5509	-107.5479	94
SAA	N=2000	-98.5458	60.8440	-	-	-104.3214	-104.3189	123

Table 15 SA vs SAA for EU-6

		VERIFICATION				CONSTRUCTION		
ALG.	step	obj	dev	$\bar{l}b^N$	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-3.2692	0.5659	-9.1974	-9.1929	-13.1477	-3.2371	0
	0.010	-3.2854	0.5654	-9.1828	-9.1782	-13.1154	-3.2470	0
	0.050	-3.4342	0.5608	-9.1341	-9.1268	-13.0708	-3.3265	1
	0.100	-3.6279	0.5575	-9.0201	-9.0084	-12.9881	-3.4259	1
	0.500	-4.8157	0.5394	-7.7299	-7.7046	-11.6856	-4.1219	0
	1.000	-5.5010	0.6512	-7.3008	-7.2726	-10.4938	-4.6646	0
	10.000	-5.8002	1.2173	-7.3728	-7.3052	-8.9175	-5.5545	1
N-SA (N=1000)	0.005	-3.3121	0.5647	-9.2181	-9.2067	-9.9742	-3.2526	3
	0.010	-3.3715	0.5632	-9.1979	-9.1864	-9.9567	-3.2827	3
	0.050	-3.8487	0.5479	-8.8674	-8.8565	-9.8440	-3.5257	3
	0.100	-4.3741	0.5163	-8.1960	-8.1924	-9.5077	-3.8138	2
	0.500	-5.9203	0.5450	-6.9083	-6.9083	-8.0532	-5.0557	3
	1.000	-6.1874	0.5714	-6.7434	-6.7118	-7.3578	-5.5638	3
	10.000	-6.2676	0.7185	-6.7578	-6.6102	-6.7471	-6.0365	3
N-SA (N=2000)	0.005	-3.3373	0.5641	-9.2038	-9.1858	-9.5292	-3.2789	5
	0.010	-3.4221	0.5620	-9.1728	-9.1555	-9.5017	-3.3212	5
	0.050	-4.0807	0.5371	-8.5959	-8.5939	-9.2867	-3.6586	5
	0.100	-4.6990	0.4988	-7.8122	-7.8122	-8.8931	-4.0284	5
	0.500	-6.0973	0.5415	-6.7822	-6.7568	-7.5327	-5.3268	6
	1.000	-6.2570	0.5726	-6.7112	-6.6218	-7.0512	-5.7515	6
	10.000	-6.2999	0.6798	-6.7664	-6.5228	-6.6111	-6.1190	6
E-SA (N=100)	0.005	-4.0352	0.5470	-8.6247	-8.6115	-12.7305	-3.6504	0
	0.010	-4.5783	0.5240	-7.9275	-7.9102	-11.8871	-3.9938	1
	0.050	-5.4874	0.5718	-7.2348	-7.1953	-10.2022	-4.8879	0
	0.100	-5.4693	0.6366	-7.3451	-7.2166	-9.9064	-5.0562	0
	0.500	-4.6187	0.6189	-7.9389	-7.8735	-10.8348	-4.5629	1
	1.000	-4.4133	0.5711	-8.1226	-8.0784	-11.1633	-4.3515	0
	10.000	-4.2599	0.5936	-8.3883	-8.3823	-11.4331	-3.9670	0
E-SA (N=1000)	0.005	-5.0592	0.4849	-7.5206	-7.5206	-8.9054	-4.2862	1
	0.010	-5.6208	0.4926	-7.1603	-7.1603	-8.3405	-4.8291	2
	0.050	-6.1205	0.5189	-6.7480	-6.7248	-7.3575	-5.6980	2
	0.100	-6.1459	0.5583	-6.7191	-6.6559	-7.1219	-5.8464	2
	0.500	-5.6174	0.6491	-7.1390	-7.0398	-7.2743	-5.4976	1
	1.000	-4.9747	0.5283	-7.5679	-7.5189	-8.1052	-4.8908	1
	10.000	-4.3580	0.5235	-8.2042	-8.1930	-8.8407	-4.1480	2
E-SA (N=2000)	0.005	-5.3939	0.4874	-7.3006	-7.3006	-8.2738	-4.5691	3
	0.010	-5.8117	0.4944	-7.0176	-7.0176	-7.8526	-5.0944	3
	0.050	-6.1948	0.5143	-6.6925	-6.6485	-7.1033	-5.8463	3
	0.100	-6.2211	0.5543	-6.6867	-6.5849	-6.9171	-5.9714	3
	0.500	-5.8539	0.7090	-7.0041	-6.7430	-6.9879	-5.7299	3
	1.000	-5.2905	0.7075	-7.3551	-7.2946	-7.5212	-5.1628	3
	10.000	-4.4801	1.1354	-8.1429	-8.1410	-8.5007	-4.2242	3
SAA	N=100	-5.5376	2.1216	-	-	-6.8526	-6.8525	26
SAA	N=1000	-6.2782	0.7413	-	-	-6.4036	-6.4036	102
SAA	N=2000	-6.3073	0.7030	-	-	-6.3658	-6.3657	171

Table 16 SA vs SAA for EU-7

ALG.	step	VERIFICATION				CONSTRUCTION		
		obj	dev	\bar{lb}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-10.1487	0.6301	-16.7163	-16.7119	-21.0028	-10.1177	0
	0.010	-10.1512	0.6301	-16.7162	-16.7117	-21.0028	-10.1192	0
	0.050	-10.1711	0.6302	-16.7150	-16.7102	-21.0028	-10.1313	0
	0.100	-10.1961	0.6303	-16.7134	-16.7083	-21.0028	-10.1464	1
	0.500	-10.4349	0.6348	-16.7012	-16.6925	-21.0028	-10.2695	1
	1.000	-10.7510	0.6493	-16.6868	-16.6719	-21.0028	-10.4278	1
	5.000	-13.2835	1.1899	-16.6515	-16.5916	-21.0028	-11.7522	0
	10.000	-14.8802	3.0185	-16.6923	-16.6556	-21.0028	-13.0563	0
N-SA (N=1000)	0.005	-10.1553	0.6301	-16.7971	-16.7844	-17.6115	-10.1173	2
	0.010	-10.1644	0.6301	-16.7971	-16.7844	-17.6115	-10.1219	2
	0.050	-10.2378	0.6304	-16.7966	-16.7839	-17.6115	-10.1590	2
	0.100	-10.3310	0.6314	-16.7961	-16.7833	-17.6115	-10.2058	3
	0.500	-11.1321	0.6664	-16.7917	-16.7778	-17.6115	-10.5988	3
	1.000	-12.1903	0.7818	-16.7871	-16.7707	-17.6115	-11.1174	2
	5.000	-15.7924	2.3172	-16.7884	-16.7433	-17.6115	-13.9232	3
	10.000	-16.0895	3.9575	-16.7985	-16.7517	-17.6916	-14.8794	3
N-SA (N=2000)	0.005	-10.1592	0.6301	-16.7993	-16.7752	-17.1800	-10.1350	5
	0.010	-10.1722	0.6301	-16.7993	-16.7752	-17.1800	-10.1415	5
	0.050	-10.2773	0.6307	-16.7988	-16.7748	-17.1800	-10.1939	5
	0.100	-10.4117	0.6327	-16.7982	-16.7743	-17.1800	-10.2604	5
	0.500	-11.5675	0.7016	-16.7931	-16.7699	-17.1800	-10.8237	5
	1.000	-12.9902	0.9087	-16.7882	-16.7649	-17.1800	-11.5501	5
	5.000	-16.0889	2.6801	-16.7911	-16.7272	-17.1800	-14.4612	5
	10.000	-16.2514	3.5233	-16.8289	-16.7060	-17.0130	-15.2442	5
E-SA (N=100)	0.005	-10.4177	0.6344	-16.7022	-16.6946	-21.0028	-10.2623	1
	0.010	-10.7019	0.6467	-16.6896	-16.6790	-21.0028	-10.4084	0
	0.050	-12.5437	0.8717	-16.6406	-16.6173	-21.0028	-11.4639	0
	0.100	-13.3191	1.0560	-16.6219	-16.5900	-21.0028	-12.1220	1
	0.500	-12.8743	1.2036	-16.6803	-16.6111	-21.0028	-12.5344	0
	1.000	-12.2309	1.3391	-16.7054	-16.6083	-21.0028	-12.0632	0
	5.000	-11.4557	0.7573	-16.7121	-16.6892	-21.0028	-11.4739	1
	10.000	-11.3683	0.8805	-16.7440	-16.7440	-38.9462	-9.4234	0
E-SA (N=1000)	0.005	-11.0297	0.6590	-16.7928	-16.7794	-17.6115	-10.5586	1
	0.010	-11.9230	0.7390	-16.7899	-16.7761	-17.6115	-11.0045	1
	0.050	-14.5576	1.2241	-16.7834	-16.7671	-17.6115	-13.1309	2
	0.100	-15.0710	1.4747	-16.7808	-16.7621	-17.6115	-13.8965	2
	0.500	-14.9006	1.8922	-16.7831	-16.7486	-17.6115	-14.3748	1
	1.000	-14.0335	1.7091	-16.7764	-16.7400	-17.6115	-13.7961	2
	5.000	-11.8605	0.7532	-16.7853	-16.7631	-17.7664	-11.8199	2
	10.000	-11.5749	0.7139	-16.7770	-16.7745	-19.4777	-11.3529	2
E-SA (N=2000)	0.005	-11.4013	0.6847	-16.7948	-16.7716	-17.1800	-10.7573	3
	0.010	-12.6465	0.8260	-16.7925	-16.7696	-17.1800	-11.3853	3
	0.050	-14.9456	1.3482	-16.7868	-16.7642	-17.1800	-13.6209	3
	0.100	-15.3818	1.6120	-16.7841	-16.7605	-17.1800	-14.3156	3
	0.500	-15.3547	2.1793	-16.7834	-16.7375	-17.1800	-14.7953	3
	1.000	-14.6640	2.0545	-16.7838	-16.7280	-17.0917	-14.2104	3
	5.000	-12.2430	1.0414	-16.7730	-16.7523	-17.2239	-12.0750	3
	10.000	-11.8296	1.5142	-16.7796	-16.7683	-17.4557	-11.5982	3
SAA	N=100	-3.3040	37.2588	-	-	-20.2507	-20.2507	8
SAA	N=1000	-15.8481	6.9923	-	-	-17.0154	-17.0154	92
SAA	N=2000	-16.1474	5.7941	-	-	-16.7027	-16.7027	162

Table 17 SA vs SAA for EU-8

ALG.	step	VERIFICATION				CONSTRUCTION		
		obj	dev	$\bar{l}b^N$	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-36.2835	6.3012	-101.9556	-101.9114	-144.8201	-35.9718	0
	0.010	-36.3124	6.3012	-101.9538	-101.9092	-144.8201	-35.9893	1
	0.050	-36.5442	6.3016	-101.9395	-101.8916	-144.8201	-36.1298	1
	0.100	-36.8355	6.3030	-101.9215	-101.8694	-144.8201	-36.3061	0
	0.500	-39.6483	6.3649	-101.7797	-101.6844	-144.8201	-37.7425	0
	1.000	-43.3600	6.5633	-101.6175	-101.4440	-144.8201	-39.5960	1
	5.000	-71.6724	13.9475	-101.3197	-100.7563	-144.8201	-54.8455	1
	10.000	-84.9435	40.5205	-101.8959	-101.5494	-144.8201	-68.2916	0
N-SA (N=1000)	0.005	-36.3602	6.3012	-102.7634	-102.6365	-110.9079	-35.9730	2
	0.010	-36.4661	6.3013	-102.7628	-102.6358	-110.9079	-36.0266	3
	0.050	-37.3210	6.3054	-102.7580	-102.6303	-110.9079	-36.4582	3
	0.100	-38.4101	6.3184	-102.7518	-102.6230	-110.9079	-37.0048	3
	0.500	-47.8370	6.7992	-102.7009	-102.5581	-110.9079	-41.6132	2
	1.000	-60.0317	8.3413	-102.6522	-102.4779	-110.9079	-47.6696	3
	5.000	-94.2198	26.2742	-102.7009	-102.2027	-110.9079	-76.7166	3
	10.000	-96.0841	35.6518	-102.7998	-102.7069	-113.2295	-84.5569	3
N-SA (N=2000)	0.005	-36.4053	6.3013	-102.7854	-102.5447	-106.5921	-36.1526	5
	0.010	-36.5565	6.3015	-102.7847	-102.5441	-106.5921	-36.2283	5
	0.050	-37.7823	6.3096	-102.7793	-102.5397	-106.5921	-36.8390	5
	0.100	-39.3542	6.3358	-102.7722	-102.5338	-106.5921	-37.6158	5
	0.500	-52.8930	7.2748	-102.7137	-102.4821	-106.5921	-44.2204	5
	1.000	-68.7131	9.9199	-102.6639	-102.4285	-106.5921	-52.5656	5
	5.000	-96.6775	30.0392	-102.7187	-102.0942	-106.5921	-81.6017	5
	10.000	-97.3613	36.3939	-103.1348	-102.6311	-106.7958	-88.4846	6
E-SA (N=100)	0.005	-39.4169	6.3579	-101.7935	-101.7127	-144.8201	-37.6452	0
	0.010	-42.7055	6.5220	-101.6526	-101.5378	-144.8201	-39.3361	1
	0.050	-62.1922	9.0853	-101.1522	-100.9075	-144.8201	-50.8899	0
	0.100	-68.9682	10.9544	-100.9908	-100.6405	-144.8201	-57.1876	0
	0.500	-62.1416	11.8625	-101.6391	-100.8601	-144.8201	-59.3630	1
	1.000	-55.9963	13.4815	-101.8781	-100.9130	-144.8201	-54.5801	0
	5.000	-49.1046	7.5264	-101.9411	-101.7109	-144.8201	-49.1509	0
	10.000	-49.1742	13.0490	-102.2136	-102.2136	-264.0121	-43.3696	1
E-SA (N=1000)	0.005	-46.4951	6.6852	-102.7147	-102.5803	-110.9079	-41.0802	2
	0.010	-56.8137	7.7224	-102.6853	-102.5454	-110.9079	-46.2404	2
	0.050	-81.7285	12.7478	-102.6202	-102.4536	-110.9079	-67.9185	2
	0.100	-86.1748	15.3124	-102.5958	-102.4023	-110.9079	-75.0241	1
	0.500	-82.4456	18.8482	-102.6128	-102.2557	-110.9079	-77.7174	2
	1.000	-72.7705	15.2295	-102.5510	-102.2047	-110.9079	-70.8866	2
	5.000	-52.5930	7.4005	-102.6446	-102.5023	-110.9079	-52.1998	2
	10.000	-50.1320	7.1088	-102.4849	-102.4849	-129.2836	-48.5130	1
E-SA (N=2000)	0.005	-50.7816	7.0227	-102.7354	-102.5038	-106.5921	-43.3549	3
	0.010	-64.5382	8.7273	-102.7125	-102.4846	-106.5921	-50.5611	3
	0.050	-85.4090	14.0222	-102.6544	-102.4285	-106.5921	-72.6737	3
	0.100	-89.2032	16.7044	-102.6282	-102.3879	-106.5921	-79.0840	3
	0.500	-87.3661	21.9272	-102.6191	-102.2588	-106.5921	-82.0989	3
	1.000	-79.3765	19.8054	-102.5945	-102.1901	-105.7417	-75.1063	3
	5.000	-55.1786	8.9174	-102.5391	-102.4659	-106.5921	-54.2121	3
	10.000	-52.6952	15.3574	-102.5993	-102.5416	-109.3439	-50.2180	3
SAA	N=100	17.9771	352.1680	-	-	-137.9611	-137.9610	6
SAA	N=1000	-93.4177	73.2566	-	-	-105.5243	-105.5203	87
SAA	N=2000	-96.5163	61.0974	-	-	-102.2914	-102.2906	160

Table 18 SA vs SAA for EU-9

ALG.	step	VERIFICATION				CONSTRUCTION		
		obj	dev	\bar{lb}^N	lb^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-339.6061	63.0118	-996.3371	-995.8950	-1424.9808	-336.4921	1
	0.010	-339.8855	63.0119	-996.3197	-995.8738	-1424.9808	-336.6617	1
	0.050	-342.1271	63.0159	-996.1800	-995.7035	-1424.9808	-338.0205	0
	0.100	-344.9437	63.0286	-996.0070	-995.4879	-1424.9808	-339.7255	0
	0.500	-372.0744	63.6080	-994.6335	-993.7017	-1424.9808	-353.6077	1
	1.000	-407.8978	65.4583	-993.0530	-991.3779	-1424.9808	-371.5070	1
	5.000	-684.5668	134.4153	-989.9175	-984.1916	-1424.9808	-519.4333	0
	10.000	-825.7714	371.5833	-995.3352	-991.8104	-1424.9808	-653.3679	0
N-SA (N=1000)	0.005	-340.3479	63.0121	-1004.4143	-1003.1457	-1085.8590	-336.4929	2
	0.010	-341.3713	63.0133	-1004.4085	-1003.1391	-1085.8590	-337.0111	2
	0.050	-349.6371	63.0511	-1004.3618	-1003.0844	-1085.8590	-341.1842	2
	0.100	-360.1603	63.1723	-1004.3020	-1003.0151	-1085.8590	-346.4668	3
	0.500	-451.1080	67.6560	-1003.8092	-1002.3886	-1085.8590	-390.9568	3
	1.000	-569.4040	82.1386	-1003.3266	-1001.6071	-1085.8590	-449.5039	2
	5.000	-915.9993	255.1959	-1003.7318	-998.8554	-1085.8590	-738.1731	3
	10.000	-939.0872	363.5217	-1004.6796	-1001.2663	-1110.0806	-824.2031	3
N-SA (N=2000)	0.005	-340.7840	63.0125	-1004.6342	-1002.2271	-1042.7008	-338.2817	5
	0.010	-342.2456	63.0148	-1004.6273	-1002.2205	-1042.7008	-339.0132	4
	0.050	-354.0951	63.0906	-1004.5751	-1002.1788	-1042.7008	-344.9174	5
	0.100	-369.2778	63.3355	-1004.5067	-1002.1223	-1042.7008	-352.4237	5
	0.500	-500.0473	72.1099	-1003.9385	-1001.6221	-1042.7008	-416.1931	5
	1.000	-654.8810	97.2201	-1003.4420	-1001.0944	-1042.7008	-497.1990	6
	5.000	-941.6887	291.9360	-1003.9342	-997.7147	-1042.7008	-788.1257	5
	10.000	-953.9882	383.8223	-1008.5464	-997.7217	-1029.0530	-858.7257	5
E-SA (N=100)	0.005	-370.7854	63.5736	-994.7223	-993.9163	-1424.9808	-353.1477	0
	0.010	-403.5081	65.1983	-993.3187	-992.1737	-1424.9808	-369.9728	0
	0.050	-598.0518	90.7247	-988.3173	-985.8744	-1424.9808	-485.1822	1
	0.100	-666.1747	109.4112	-986.6931	-983.2019	-1424.9808	-548.2724	0
	0.500	-598.7146	118.6970	-993.1622	-985.3957	-1424.9808	-570.7319	0
	1.000	-537.1122	134.7827	-995.5442	-985.8843	-1424.9808	-522.8490	1
	5.000	-467.9027	75.2778	-996.1817	-993.8806	-1424.9808	-468.3696	0
	10.000	-468.5466	130.5814	-998.9136	-998.9137	-2618.0620	-411.5473	0
E-SA (N=1000)	0.005	-441.2167	66.8149	-1003.9286	-1002.5858	-1085.8590	-387.3251	2
	0.010	-543.9050	77.0978	-1003.6355	-1002.2361	-1085.8590	-438.6718	1
	0.050	-793.6082	127.2962	-1002.9834	-1001.3190	-1085.8590	-655.3626	2
	0.100	-838.3218	152.9272	-1002.7401	-1000.8066	-1085.8590	-726.6114	2
	0.500	-801.7337	188.5334	-1002.9137	-999.3462	-1085.8590	-754.2690	2
	1.000	-705.3121	152.7011	-1002.2903	-998.8188	-1085.8590	-686.3337	1
	5.000	-502.9805	74.0455	-1003.2255	-1001.8013	-1085.8590	-499.0478	1
	10.000	-480.5189	126.7385	-1001.5706	-1001.5706	-1202.6516	-462.7362	2
E-SA (N=2000)	0.005	-483.8734	70.1595	-1004.1355	-1001.8192	-1042.7008	-409.9674	3
	0.010	-621.0829	87.1127	-1003.9064	-1001.6274	-1042.7008	-481.7079	3
	0.050	-830.5069	140.0433	-1003.3258	-1001.0670	-1042.7008	-702.9633	3
	0.100	-868.6242	166.8363	-1003.0636	-1000.6626	-1042.7008	-767.2545	3
	0.500	-850.8188	219.2586	-1002.9677	-999.3737	-1042.7008	-798.0241	3
	1.000	-771.2904	198.3226	-1002.7299	-998.6815	-1034.2088	-728.4992	3
	5.000	-529.0234	89.3518	-1002.1681	-1001.4363	-1042.7008	-519.2999	4
	10.000	-503.8182	153.3248	-1002.7689	-1002.1950	-1070.2659	-479.2331	3
SAA	N=100	201.2264	3521.2462	-	-	-1356.8687	-1356.8677	7
SAA	N=1000	-910.8297	734.8461	-	-	-1032.3006	-1032.2738	91
SAA	N=2000	-941.9854	611.0414	-	-	-999.9114	-999.8982	161

Table 19 SA vs SAA for EU-10

		VERIFICATION				CONSTRUCTION		
ALG.	step	obj	dev	\bar{b}^N	b^N	\bar{f}^N	ub	time
N-SA (N=100)	0.005	-1676.8591	315.0591	-4960.2040	-4957.9885	-7103.4410	-1661.1747	0
	0.010	-1678.5487	315.0600	-4960.0990	-4957.8610	-7103.4410	-1662.1996	0
	0.050	-1692.1085	315.0894	-4959.2590	-4956.8288	-7103.4410	-1670.4180	1
	0.100	-1709.1651	315.1817	-4958.2084	-4955.5264	-7103.4410	-1680.7371	1
	0.500	-1875.6770	319.4277	-4949.9686	-4944.6357	-7103.4410	-1765.0413	0
	1.000	-2094.9050	333.1882	-4940.8992	-4930.5815	-7103.4410	-1874.2415	0
	5.000	-3650.9872	840.3725	-4931.2576	-4905.6257	-7103.4410	-2747.9567	0
	10.000	-3800.7619	1630.9513	-4977.4536	-4977.4725	-13023.9623	-2090.9446	1
N-SA (N=1000)	0.005	-1681.3447	315.0616	-5000.6024	-4994.2585	-5407.8319	-1661.5425	3
	0.010	-1687.5350	315.0701	-5000.5675	-4994.2185	-5407.8319	-1664.6766	3
	0.050	-1737.6309	315.3476	-5000.2840	-4993.8904	-5407.8319	-1689.9472	3
	0.100	-1801.6284	316.2411	-4999.9201	-4993.4581	-5407.8319	-1722.0148	2
	0.500	-2358.9495	349.5722	-4996.9862	-4989.6317	-5407.8319	-1993.8515	3
	1.000	-3055.9686	452.0546	-4994.5041	-4985.1447	-5407.8319	-2347.5158	3
	5.000	-4632.4882	1521.6509	-4998.8894	-4971.5507	-5407.8319	-3826.7989	2
	10.000	-4647.9744	1923.8632	-5001.4221	-4999.3838	-5761.7873	-4164.6643	3
N-SA (N=2000)	0.005	-1683.9821	315.0644	-5001.7014	-4989.6669	-5192.0409	-1670.7092	5
	0.010	-1692.8258	315.0813	-5001.6535	-4989.6352	-5192.0409	-1675.1336	5
	0.050	-1764.7102	315.6381	-5001.3410	-4989.3671	-5192.0409	-1710.9096	5
	0.100	-1857.2203	317.4494	-5000.9237	-4989.0187	-5192.0409	-1756.5379	5
	0.500	-2652.4518	382.0028	-4997.5860	-4986.0201	-5192.0409	-2145.5733	5
	1.000	-3512.5867	546.2880	-4995.1150	-4982.9658	-5192.0409	-2621.5074	6
	5.000	-4729.8534	1746.7144	-4999.2312	-4967.9144	-5192.0409	-4045.2517	5
	10.000	-4713.8299	1666.9798	-5015.0647	-5015.0645	-5282.8438	-4313.8041	5
E-SA (N=100)	0.005	-1832.4637	317.8680	-4952.1495	-4948.1207	-7103.4410	-1744.2755	0
	0.010	-1996.0775	325.9916	-4945.1242	-4939.4039	-7103.4410	-1828.4007	0
	0.050	-2968.7958	453.6234	-4920.1242	-4907.9057	-7103.4410	-2404.4480	1
	0.100	-3309.4105	547.0562	-4912.0031	-4894.5463	-7103.4410	-2719.8988	0
	0.500	-2972.1098	593.4851	-4944.3479	-4905.5160	-7103.4410	-2832.1963	0
	1.000	-2664.0980	673.9133	-4956.2551	-4907.9598	-7103.4410	-2592.7819	1
	5.000	-2318.0505	376.3888	-4959.4452	-4947.9391	-7103.4410	-2320.3848	0
	10.000	-2321.2721	652.8988	-4973.0991	-4973.1006	-13069.9200	-2037.3464	0
E-SA (N=1000)	0.005	-2184.6204	334.0744	-4998.1803	-4991.4658	-5407.8319	-1915.1624	2
	0.010	-2698.0620	385.4891	-4996.7147	-4989.7212	-5407.8319	-2171.8957	2
	0.050	-3946.5778	636.4811	-4993.4538	-4985.1330	-5407.8319	-3255.3498	1
	0.100	-4170.1457	764.6360	-4992.2371	-4982.5620	-5407.8319	-3611.5939	2
	0.500	-3987.2053	942.6669	-4993.1057	-4975.2666	-5407.8319	-3749.8819	2
	1.000	-3505.0974	763.5056	-4989.9914	-4972.6317	-5407.8319	-3410.2052	1
	5.000	-2493.4395	370.2274	-4994.6645	-4987.5433	-5407.8319	-2473.7761	2
	10.000	-2381.1315	633.6926	-4984.7356	-4984.7356	-6029.5468	-2289.8039	2
E-SA (N=2000)	0.005	-2397.9041	350.7973	-4999.2141	-4987.6338	-5192.0409	-2028.3740	4
	0.010	-3083.9515	435.5634	-4998.0716	-4986.6741	-5192.0409	-2387.0764	3
	0.050	-4131.0714	700.2164	-4995.1679	-4983.8716	-5192.0409	-3493.3532	3
	0.100	-4321.6580	834.1813	-4993.8571	-4981.8515	-5192.0409	-3814.8095	3
	0.500	-4232.6307	1096.2932	-4993.3757	-4975.4059	-5192.0409	-3968.6575	4
	1.000	-3834.9890	991.6129	-4992.1858	-4971.9463	-5149.5918	-3621.0437	3
	5.000	-2623.6539	446.7588	-4989.3751	-4985.7188	-5192.0409	-2575.0362	3
	10.000	-2497.6344	766.5842	-4992.3817	-4989.5130	-5329.9094	-2374.7451	3
SAA	N=100	1039.0135	17631.6927	-	-	-6763.1581	-6763.1456	7
SAA	N=1000	-4530.8206	3687.4747	-	-	-5140.1782	-5140.1360	89
SAA	N=2000	-4688.9239	3053.7409	-	-	-4978.2333	-4978.1967	161