A Comment on "Computational Complexity of Stochastic Programming Problems"

Grani A. Hanasusanto[1], Daniel Kuhn[1], and Wolfram Wiesemann[2]

[1]*Risk Analytics and Optimization Chair, École Polytechnique Fédérale de Lausanne, Switzerland*

[2]*Imperial College Business School, Imperial College London, United Kingdom*

October 6, 2015

Abstract

Although stochastic programming problems were always believed to be computationally challenging, this perception has only recently received a theoretical justification by the seminal work of Dyer and Stougie (*Mathematical Programming A*, 106(3):423–432, 2006). Amongst others, that paper argues that linear two-stage stochastic programs with fixed recourse are #P-hard even if the random problem data is governed by independent uniform distributions. We show that Dyer and Stougie's proof is not correct, and we offer a correction which establishes the stronger result that even the approximate solution of such problems is #P-hard for a sufficiently high accuracy. We also provide new results which indicate that linear two-stage stochastic programs with random recourse seem even more challenging to solve.

Keywords: stochastic programming; complexity theory.

1 Introduction

We consider linear two-stage stochastic programs of the form

$$
\begin{aligned}
\text{maximize} \quad & \boldsymbol{c}^\top \boldsymbol{x} + \mathbb{E}\left[Q(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})\right] \\
\text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \ \ \boldsymbol{x} \in \mathbb{R}^{n_1}_+,
\end{aligned}
\tag{1a}
$$

where the problem data $\boldsymbol{c} \in \mathbb{R}^{n_1}$, $\boldsymbol{A} \in \mathbb{R}^{m_1 \times n_1}$ and $\boldsymbol{b} \in \mathbb{R}^{m_1}$ is deterministic. For a fixed realization $\boldsymbol{\xi} \in \mathbb{R}^k$ of the random vector $\tilde{\boldsymbol{\xi}}$, $Q(\boldsymbol{x}, \boldsymbol{\xi})$ denotes the optimal value of the second-stage problem

$$
\begin{aligned}
\text{maximize} \quad & \boldsymbol{q}(\boldsymbol{\xi})^\top \boldsymbol{y} \\
\text{subject to} \quad & \boldsymbol{T}(\boldsymbol{\xi})\boldsymbol{x} + \boldsymbol{W}(\boldsymbol{\xi})\boldsymbol{y} = \boldsymbol{h}(\boldsymbol{\xi}), \ \ \boldsymbol{y} \in \mathbb{R}_+^{n_2},
\end{aligned}
\tag{1b}
$$

where $\boldsymbol{q} : \mathbb{R}^k \to \mathbb{R}^{n_2}$, $\boldsymbol{T} : \mathbb{R}^k \to \mathbb{R}^{m_2 \times n_1}$, $\boldsymbol{W} : \mathbb{R}^k \to \mathbb{R}^{m_2 \times n_2}$ and $\boldsymbol{h} : \mathbb{R}^k \to \mathbb{R}^{m_2}$ depend affinely on $\boldsymbol{\xi}$. The stochastic program (1) has *fixed recourse* if the recourse matrix $\boldsymbol{W}$ does not depend on the realization of $\tilde{\boldsymbol{\xi}}$, and it displays *random recourse* otherwise. By convention, the optimal values of (1a) and (1b) are defined to be $+\infty$ if the respective problems are infeasible. For a survey of the vast stochastic programming literature, which dates back to the 1950s, we refer to [2, 10, 11].

While stochastic programming problems were always believed to be difficult to solve, this suspicion has only been confirmed recently by the seminal work of Dyer and Stougie [4]. Amongst others, that paper argues that calculating the expected recourse value $\mathbb{E}[Q(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})]$ for a fixed first-stage decision $\boldsymbol{x}$ is #P-hard[1] even if the stochastic program (1) exhibits fixed recourse and the random vector $\tilde{\boldsymbol{\xi}}$ is governed by a uniform distribution supported on $[0, 1]^k$. This result has attracted significant attention in the stochastic programming and robust optimization communities, where it provides a formal justification for developing approximate solutions schemes. The proof of this statement studies the expected recourse value $\mathcal{Q}(\boldsymbol{\alpha}, \beta) = \mathbb{E}[Q(\tilde{\boldsymbol{\xi}}; \boldsymbol{\alpha}, \beta)]$ of the second-stage problem

$$
\begin{aligned}
\text{maximize} \quad & \boldsymbol{\xi}^\top \boldsymbol{y} - \beta z \\
\text{subject to} \quad & \boldsymbol{y} \le \boldsymbol{\alpha} z, \ \ \boldsymbol{y} \in \mathbb{R}_+^k, \ \ z \in [0, 1]
\end{aligned}
\tag{2}
$$

under the uniform distribution supported on $[0, 1]^k$, where $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$ are fixed parameters. Note that problem (2) does not involve any first-stage decisions. It is claimed in [4] that $1 - \mathcal{Q}(\boldsymbol{\alpha}, \beta)$ coincides with the volume of the knapsack polytope $P(\boldsymbol{\alpha}, \beta) = \{\boldsymbol{\xi} \in [0, 1]^k : \boldsymbol{\alpha}^\top \boldsymbol{\xi} \le \beta\}$, the calculation of which is known to be #P-hard [5]. We now show that this claim is false.

Proposition 1. *For $\boldsymbol{\alpha} = \mathbf{e}$ and $\beta = 1$, where $\mathbf{e} \in \mathbb{R}^k$ is the vector of all ones, the expected recourse*

---

[1]The complexity class #P contains the counting problems associated with the decision problems in the complexity class NP (*e.g.*, counting the number of Hamiltonian cycles in a graph), see [6, 9]. Thus, a counting problem is in #P if the items to be counted (*e.g.*, the Hamiltonian cycles) can be validated as such in polynomial time. By definition, a #P problem is at least as difficult as the corresponding NP problem. It is therefore commonly believed that #P-hard problems, which are the hardest problems in #P, do not admit polynomial-time solution methods.

value of problem (2) *satisfies*

$$\mathcal{Q}(\boldsymbol{\alpha}, \beta) = \frac{k-2}{2} + \frac{1}{(k+1)!}.$$

Proof. It is shown in [4] that for a given realization $\boldsymbol{\xi}$ of the random vector $\tilde{\boldsymbol{\xi}}$, the unique optimal second-stage decisions are $\boldsymbol{y} = \boldsymbol{0}$ and $z = 0$ if $\boldsymbol{\alpha}^\top \boldsymbol{\xi} \leq \beta$ and $\boldsymbol{y} = \boldsymbol{\alpha}$ and $z = 1$ otherwise. We thus conclude that the expected recourse value of problem (2) satisfies

$$
\begin{aligned}
\mathcal{Q}(\boldsymbol{\alpha}, \beta) &= \mathbb{E}\left[\max\left\{\mathbf{e}^\top \tilde{\boldsymbol{\xi}} - 1, 0\right\}\right] \\
&= \mathbb{E}\left[\mathbf{e}^\top \tilde{\boldsymbol{\xi}} - 1\right] + \mathbb{E}\left[\max\left\{1 - \mathbf{e}^\top \tilde{\boldsymbol{\xi}}, 0\right\}\right] \\
&= \frac{k-2}{2} + \int_0^1 \mathrm{d}\xi_1 \int_0^{1-\xi_1} \mathrm{d}\xi_2 \cdots \int_0^{1-\sum_{i=1}^{k-1}\xi_i} \mathrm{d}\xi_k \left(1 - \mathbf{e}^\top \boldsymbol{\xi}\right) \\
&= \frac{k-2}{2} + \frac{1}{(k+1)!},
\end{aligned}
$$

where the integral in the penultimate line can be viewed as the normalization constant of a Dirichlet distribution of order $k+1$ with parameters $(\mathbf{e}^\top, 2)^\top$ and thus evaluates to $1/(k+1)!$, see [1]. □

Proposition 1 implies that $1 - \mathcal{Q}(\mathbf{e}, 1)$ is strictly negative for $k \geq 4$ and can therefore not be equal to the volume of the polytope $P(\mathbf{e}, 1)$.

The remainder of this paper develops as follows. In Section 2 we correct the proof of Dyer and Stougie and show that calculating the expected recourse value of a linear two-stage stochastic program with *fixed recourse* is indeed #P-hard. Moreover, we strengthen the original result by showing that approximating this expected recourse value within an exponentially small error remains #P-hard. Similar arguments as in [4] then allow us to conclude that finding an approximately optimal solution is #P-hard as well. Section 3 shows that linear two-stage stochastic program with *random recourse* are #P-hard to approximate within an exponentially small error, and NP-hard to approximate within an inverse polynomial error. Unless P = NP, these stronger results preclude the existence of fully polynomial-time approximation schemes (FPTAS) for the generic linear two-stage stochastic program (1), see [6]. We provide concluding remarks in Section 4.

Notation Vectors and matrices are denoted by bold lower-case and upper-case letters, respectively, while scalars are printed in regular font. We notationally highlight random objects by a tilde sign. We use $\mathbf{e}_i$ and $\mathbf{e}$ to refer to the i-th canonical basis vector and the vector of ones, respectively. In both cases, the dimension will be clear from the context. We define the indicator function $\mathbb{I}_{[\mathcal{E}]}$

as $\mathbb{I}_{[\mathcal{E}]} = 1$ if the logical expression $\mathcal{E}$ holds true and 0 otherwise. For any matrix $\boldsymbol{M} \in \mathbb{R}^{m \times n}$, we define its norm by $\|\boldsymbol{M}\|_1 = \max_{\boldsymbol{x} \in \mathbb{R}^n} \{\|\boldsymbol{M}\boldsymbol{x}\|_1 \; : \; \|\boldsymbol{x}\|_1 = 1\}$.

2 Stochastic Programs with Fixed Recourse

In this section we correct the #P-hardness proof of Dyer and Stougie for linear two-stage stochastic programs with fixed recourse. To this end, we show first that calculating the volume of a knapsack polytope $P(\boldsymbol{\alpha}, \beta) = \{\boldsymbol{\xi} \in [0, 1]^k \; : \; \boldsymbol{\alpha}^\top \boldsymbol{\xi} \leq \beta\}$ to within a sufficiently high accuracy is #P-hard.

Lemma 1. *Computing the volume of the knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ for $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$ to within an absolute accuracy of ϵ is #P-hard whenever*

$$\epsilon < \frac{1}{2k!(\|\boldsymbol{\alpha}\|_1 + 2)^k (k+1)^{k+1} \prod_{j=1}^k \alpha_j}. \tag{3}$$

The proof of Lemma 1 relies on a reduction from the #PARITY problem, which is known to be #P-hard [5, Lemma 1].

#PARITY

Instance. Given are $\boldsymbol{\alpha} \in \mathbb{N}^k$ and $\beta \in \mathbb{N}$, which encode the knapsack polytope $P(\boldsymbol{\alpha}, \beta)$.

Question. For $K = P(\boldsymbol{\alpha}, \beta) \cap \{0, 1\}^k$, compute

$$D = |\{\boldsymbol{\xi} \in K \; : \; \mathbf{e}^\top \boldsymbol{\xi} \text{ is even}\}| - |\{\boldsymbol{\xi} \in K \; : \; \mathbf{e}^\top \boldsymbol{\xi} \text{ is odd}\}|.$$

Proof of Lemma 1. Consider an instance of the #PARITY problem with input $\boldsymbol{\alpha} \in \mathbb{N}^k$ and $\beta \in \mathbb{N}$, and select any non-negative ϵ satisfying (3). If $\boldsymbol{\alpha}^\top \mathbf{e} < \beta$, then the knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ reduces to the unit hypercube, in which case the solution to the #PARITY problem is $D = 0$. From now on we may thus focus on the case $\boldsymbol{\alpha}^\top \mathbf{e} \geq \beta$.

Introduce knapsack budgets $\gamma_j = (\beta + \frac{j}{k+1})$, $j = 0, \ldots, k$, and define the Vandermonde matrix

$$\boldsymbol{F} = \begin{bmatrix} (\gamma_0)^k & (\gamma_0)^{k-1} & \cdots & (\gamma_0)^0 \\ \vdots & \vdots & \ddots & \vdots \\ (\gamma_k)^k & (\gamma_k)^{k-1} & \cdots & (\gamma_k)^0 \end{bmatrix}.$$

Note that

$$\det(\boldsymbol{F}) \;=\; \prod_{0 \leq i < j \leq k} (\gamma_j - \gamma_i) \;\neq\; 0$$

4

because all knapsack budgets are mutually distinct. Thus, $\boldsymbol{F}$ is invertible. We also find

$$\|\boldsymbol{F}^{-1}\|_1 = \max_{i=0,\ldots,k} \prod_{\substack{j=0 \\ j \neq i}}^{k} \frac{1 + |\gamma_j|}{|\gamma_j - \gamma_i|} \leq \frac{(2+\beta)^k}{(\frac{1}{k+1})^k} \leq (\|\boldsymbol{\alpha}\|_1 + 2)^k (k+1)^k, \tag{4}$$

where the first equality follows from [7, Theorem 1] and the observation that $\|\boldsymbol{F}^{-1}\|_1 = \|(\boldsymbol{F}^{-1})^\top\|_\infty$, which holds since the 1-norm of a matrix is the maximum of its column sums and the ∞-norm is the maximum of its row sums. The last inequality follows from the fact that $\boldsymbol{\alpha}^\top \mathbf{e} \geq \beta$. Next, set $\boldsymbol{g} = [\mathrm{Vol}(P(\boldsymbol{\alpha}, \gamma_0)), \ldots, \mathrm{Vol}(P(\boldsymbol{\alpha}, \gamma_k))]^\top$, which contains the volumes of the $k+1$ knapsack polytopes with budgets $\gamma_0, \ldots, \gamma_k$. As $\boldsymbol{F}$ is invertible, the system of linear equations

$$\boldsymbol{F}\boldsymbol{x} = \left(k! \prod_{j=1}^{k} \alpha_j \right) \boldsymbol{g}$$

has a unique solution $\boldsymbol{x}^\star \in \mathbb{R}^{k+1}$. We know from the proof of Theorem 1 in [5] that $\boldsymbol{x}^\star$ is in fact an integer vector and that its first element coincides with the solution D of the #PARITY problem.

In what follows we assume that the vector $\boldsymbol{g}$ of the different knapsack polytopes' exact volumes is not available but that we can compute an approximation $\boldsymbol{g}_\epsilon = \boldsymbol{g} + \epsilon \boldsymbol{d}$ with $\boldsymbol{d} \in [-1, 1]^{k+1}$. Let $\boldsymbol{x}_\epsilon^\star$ be the solution of the perturbed system

$$\boldsymbol{F}\boldsymbol{x} = \left(k! \prod_{j=1}^{k} \alpha_j \right) \boldsymbol{g}_\epsilon, \tag{5}$$

and note that $\boldsymbol{x}_\epsilon^\star$ can be computed efficiently by Gaussian elimination, for instance. By left-multiplying the above equation with $\boldsymbol{F}^{-1}$ we obtain

$$\boldsymbol{x}_\epsilon^\star = \left(k! \prod_{j=1}^{k} \alpha_j \right) \boldsymbol{F}^{-1} \boldsymbol{g} + \epsilon \left(k! \prod_{j=1}^{k} \alpha_j \right) \boldsymbol{F}^{-1} \boldsymbol{d} = \boldsymbol{x}^\star + \epsilon \left(k! \prod_{j=1}^{k} \alpha_j \right) \boldsymbol{F}^{-1} \boldsymbol{d},$$

where the second equality follows from the definition of $\boldsymbol{x}^\star$. Thus, we find

$$\|\boldsymbol{x}_\epsilon^\star - \boldsymbol{x}^\star\|_1 = \epsilon \left(k! \prod_{j=1}^{k} \alpha_j \right) \|\boldsymbol{F}^{-1} \boldsymbol{d}\|_1 < \frac{1}{2} \frac{\|\boldsymbol{F}^{-1}\|_1 \|\boldsymbol{d}\|_1}{(\|\boldsymbol{\alpha}\|_1 + 2)^k (1+k)^{k+1}} \leq \frac{\|\boldsymbol{d}\|_1}{2(k+1)} \leq \frac{1}{2}, \tag{6}$$

where the first inequality holds due to our choice of ϵ and the definition of the matrix norm, the second inequality follows from (4), and the third inequality holds because $\boldsymbol{d} \in [-1, 1]^{k+1}$. As any two corresponding components of $\boldsymbol{x}_\epsilon^\star$ and $\boldsymbol{x}^\star$ differ by strictly less than $\frac{1}{2}$ and as the first component of $\boldsymbol{x}^\star$ coincides with D, we can compute D by rounding the first component of $\boldsymbol{x}_\epsilon^\star$ to the nearest

integer. In summary, we have found the following procedure for solving the #PARITY problem: (i) construct $\boldsymbol{F}$, (ii) determine $\boldsymbol{g}_\epsilon$ by calling $k+1$ times an algorithm for computing the volume of a knapsack polytope to within accuracy ϵ, (iii) solve the system of linear equations (5) to obtain $\boldsymbol{x}_\epsilon^\star$, and (iv) round the first component of $\boldsymbol{x}_\epsilon^\star$ to the nearest integer. All operations with the exception of the volume calculations can be carried out in time polynomial in the bit length of $\boldsymbol{\alpha}$ and β. Thus, if we could approximate the volume of $P(\boldsymbol{\alpha}, \beta)$ for $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$ in time polynomial in the bit length of $\boldsymbol{\alpha}$ and β, then we could solve the #PARITY problem efficiently. We conclude that approximating the volume of a knapsack polytope is at least as hard as solving the #P-hard #PARITY problem. $\qquad\square$

In the remainder we denote by $\mathcal{Q}(\boldsymbol{\alpha}, \beta) = \mathbb{E}[Q(\tilde{\boldsymbol{\xi}}; \boldsymbol{\alpha}, \beta)]$ the expected recourse value of the second-stage problem (2). We first show that the volume of the knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ can be expressed in terms of the partial derivative $\partial\mathcal{Q}(\boldsymbol{\alpha}, \beta)/\partial\beta$, which we abbreviate as $\mathcal{Q}'(\boldsymbol{\alpha}, \beta)$.

Lemma 2. *For every $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$ we have $\mathrm{Vol}(P(\boldsymbol{\alpha}, \beta)) = \mathcal{Q}'(\boldsymbol{\alpha}, \beta) + 1$.*

Proof. Recall that the unique optimal second-stage decisions of problem (2) are $\boldsymbol{y} = \boldsymbol{0}$ and $z = 0$ if $\boldsymbol{\alpha}^\top \boldsymbol{\xi} \le \beta$ and $\boldsymbol{y} = \boldsymbol{\alpha}$ and $z = 1$ otherwise, see [4]. Thus, we have $\mathcal{Q}(\boldsymbol{\alpha}, \beta) = \mathbb{E}[\max\{\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\xi}} - \beta, 0\}]$. As $\mathbb{E}[\tilde{\boldsymbol{\xi}}] = \mathbf{e}/2$ and because the relation

$$\max\{\gamma - \beta, 0\} = \gamma - \int_0^\beta \mathbb{I}_{[\gamma \ge y]} \, \mathrm{d}y$$

holds for any $\beta \in \mathbb{R}$ and $\gamma \in \mathbb{R}_+$, we find

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\alpha}, \beta) &= \frac{\boldsymbol{\alpha}^\top\mathbf{e}}{2} - \mathbb{E}\left[\int_0^\beta \mathbb{I}_{[\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\xi}} \ge y]} \, \mathrm{d}y\right] \\ &= \frac{\boldsymbol{\alpha}^\top\mathbf{e}}{2} - \int_0^\beta \mathbb{E}\left[\mathbb{I}_{[\boldsymbol{\alpha}^\top\tilde{\boldsymbol{\xi}} \ge y]}\right] \, \mathrm{d}y \\ &= \frac{\boldsymbol{\alpha}^\top\mathbf{e}}{2} - \int_0^\beta [1 - \mathrm{Vol}(P(\boldsymbol{\alpha}, y))] \, \mathrm{d}y, \end{aligned}$$

where the second equality follows from Fubini's theorem. Note that the volume $\mathrm{Vol}(P(\boldsymbol{\alpha}, \beta))$ of the knapsack polytope changes continuously with $\beta \in \mathbb{R}$, and thus $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ is continuously differentiable in β. The claim then follows by differentiating both sides of the above relation with respect to β. $\quad\square$

We are now armed to prove the #P-hardness of computing $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ to high accuracy. To this end, we show that calculating $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ to high accuracy allows us to closely approximate

its derivative $\mathcal{Q}'(\boldsymbol{\alpha}, \beta)$ and thus, by virtue of Lemma 2, the volume of the knapsack polytope $\mathrm{Vol}(P(\boldsymbol{\alpha}, \beta))$. The #P-hardness then follows from Lemma 1.

Theorem 1. *Computing the expected recourse value $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ of the second-stage problem* (2) *for $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$ to within an absolute accuracy of δ is #P-hard whenever*

$$\delta < \left(\frac{\alpha_k}{2k!(1 + \alpha_k)(\|\boldsymbol{\alpha}\|_1 + 2)^k (k+1)^{k+1} \prod_{j=1}^k \alpha_j} \right)^2. \tag{7}$$

Proof. Consider a knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ encoded by the parameters $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$, where we assume without loss of generality that $k \geq 2$. Select any non-negative δ satisfying (7) and set $h = 2\sqrt{\delta}$. In the following, we denote by $\mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta)$ and $\mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta + h)$ any δ-approximations of the expected recourse values $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ and $\mathcal{Q}(\boldsymbol{\alpha}, \beta + h)$, respectively. Then, we obtain

$$\left| \frac{\mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta + h) - \mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta)}{h} + 1 - \mathrm{Vol}(P(\boldsymbol{\alpha}, \beta)) \right|$$

$$= \left| \frac{\mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta + h) - \mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta)}{h} - \mathcal{Q}'(\boldsymbol{\alpha}, \beta) \right|$$

$$\leq \left| \frac{\mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta + h) - \mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta)}{h} - \frac{\mathcal{Q}(\boldsymbol{\alpha}, \beta + h) - \mathcal{Q}(\boldsymbol{\alpha}, \beta)}{h} \right| + \left| \frac{\mathcal{Q}(\boldsymbol{\alpha}, \beta + h) - \mathcal{Q}(\boldsymbol{\alpha}, \beta)}{h} - \mathcal{Q}'(\boldsymbol{\alpha}, \beta) \right|$$

$$\leq \frac{2\delta}{h} + \frac{h}{2} \sup_{\ell \in [0,h]} \left| \mathcal{Q}''(\boldsymbol{\alpha}, \ell) \right| \;=\; \sqrt{\delta} \left(1 + \sup_{\ell \in [0,h]} \left| \mathcal{Q}''(\boldsymbol{\alpha}, \ell) \right| \right),$$

where the first equality holds due to Lemma 2, and the second inequality follows from Taylor's theorem. To derive an upper bound on the last expression, we observe that $\mathcal{Q}''(\boldsymbol{\alpha}, \beta) = \partial \mathrm{Vol}(P(\boldsymbol{\alpha}, \beta))/\partial\beta \geq 0$ by virtue of Lemma 2. As shown in [5], the volume of the knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ admits the following closed-form expression.

$$\mathrm{Vol}(P(\boldsymbol{\alpha}, \beta)) = \frac{\sum_{\boldsymbol{v} \in \{0,1\}^k} (-1)^{\mathbf{e}^\top \boldsymbol{v}} \max\{0, \beta - \boldsymbol{\alpha}^\top \boldsymbol{v}\}^k}{k! \prod_{j=1}^k \alpha_j} \tag{8}$$

To simplify the following derivation, we introduce the shorthand notation $\boldsymbol{\alpha}_{k-1} = (\alpha_1, \ldots, \alpha_{k-1})^\top$.

Differentiating the right-hand side of (8) with respect to β yields

$$
\begin{aligned}
&\mathcal{Q}''(\boldsymbol{\alpha}, \beta) \\[4pt]
&= \frac{\sum_{\boldsymbol{v} \in \{0,1\}^k} (-1)^{\mathbf{e}^\top \boldsymbol{v}} \max\{0, \beta - \boldsymbol{\alpha}^\top \boldsymbol{v}\}^{k-1}}{(k-1)! \prod_{j=1}^k \alpha_j} \\[4pt]
&= \frac{\sum_{\boldsymbol{v} \in \{0,1\}^k : v_k = 0} (-1)^{\mathbf{e}^\top \boldsymbol{v}} \max\{0, \beta - \boldsymbol{\alpha}^\top \boldsymbol{v}\}^{k-1} + \sum_{\boldsymbol{v} \in \{0,1\}^k : v_k = 1} (-1)^{\mathbf{e}^\top \boldsymbol{v}} \max\{0, \beta - \boldsymbol{\alpha}^\top \boldsymbol{v}\}^{k-1}}{(k-1)! \prod_{j=1}^k \alpha_j} \\[4pt]
&= \frac{\sum_{\boldsymbol{v} \in \{0,1\}^{k-1}} (-1)^{\mathbf{e}^\top \boldsymbol{v}} \max\{0, \beta - \boldsymbol{\alpha}_{k-1}^\top \boldsymbol{v}\}^{k-1} - \sum_{\boldsymbol{v} \in \{0,1\}^{k-1}} (-1)^{\mathbf{e}^\top \boldsymbol{v}} \max\{0, \beta - \boldsymbol{\alpha}_{k-1}^\top \boldsymbol{v} - \alpha_k\}^{k-1}}{(k-1)! \prod_{j=1}^k \alpha_j} \\[4pt]
&= \frac{1}{\alpha_k} \left(\mathrm{Vol}(P(\boldsymbol{\alpha}_{k-1}, \beta)) - \mathrm{Vol}(P(\boldsymbol{\alpha}_{k-1}, \beta - \alpha_k)) \right) \leq \frac{1}{\alpha_k}.
\end{aligned}
$$

Since this upper bound is independent of β, we find that $\sup_{\ell \in [0,h]} \mathcal{Q}''(\boldsymbol{\alpha}, \ell) \leq \frac{1}{\alpha_k}$. In summary, we have thus shown that

$$
\left| \frac{\mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta + h) - \mathcal{Q}_\delta(\boldsymbol{\alpha}, \beta)}{h} + 1 - \mathrm{Vol}(P(\boldsymbol{\alpha}, \beta)) \right| \leq \sqrt{\delta} \left(1 + \frac{1}{\alpha_k} \right) < \epsilon,
$$

where $\epsilon = 1/(2k!(\|\boldsymbol{\alpha}\|_1 + 2)^k (k+1)^{k+1} \prod_{j=1}^k \alpha_j)$. Thus, we can compute the volume of the knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ to within an absolute accuracy of ϵ by computing the expected recourse values of $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ and $\mathcal{Q}(\boldsymbol{\alpha}, \beta + h)$ to within accuracy δ, dividing their difference by h and adding 1. With the exception of computing the expected recourse values, all of these operations can be carried out in time polynomial in the bit length of the instance of (1) that corresponds to problem (2). We conclude that computing the expected recourse value $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ to within accuracy δ is at least as hard as computing the volume of a knapsack polytope $P(\boldsymbol{\alpha}, \beta)$ to within accuracy ϵ. The claim then follows from Lemma 1. $\qquad\square$

Since the expected recourse value of problem (2) can be computed through a linear two-stage stochastic program (1) with fixed recourse, we conclude that approximating the optimal value of a linear two-stage stochastic program with fixed recourse to high accuracy is #P-hard as well.

Theorem 1 extends to the problem of determining an approximately optimal decision to problem (1). To this end, we call a feasible solution $\boldsymbol{x}$ to problem (1) ϵ-*optimal* if $|f^\star - (\boldsymbol{c}^\top \boldsymbol{x} + \mathbb{E}[Q(\boldsymbol{x}, \tilde{\boldsymbol{\xi}})])| / \max\{|f^\star|, 1\} \leq \epsilon$, where $f^\star$ denotes the optimal value of problem (1).

For fixed $c \in \mathbb{R}$, we consider the following instance of problem (1) with fixed recourse:

$$
\begin{aligned}
\text{maximize} \quad & -cx + \mathbb{E}\left[Q(x, \tilde{\boldsymbol{\xi}}; \boldsymbol{\alpha}, \beta) \right] \\
\text{subject to} \quad & x \in [0, 1],
\end{aligned}
\tag{9}
$$

8

where $\tilde{\boldsymbol{\xi}}$ follows a uniform distribution supported on $[0,1]^k$, and for a fixed realization $\boldsymbol{\xi}$ the recourse function $Q(x, \boldsymbol{\xi}; \boldsymbol{\alpha}, \beta)$ is the optimal value of the second-stage problem

$$
\begin{aligned}
\text{maximize} \quad & \boldsymbol{\xi}^\top \boldsymbol{y} - \beta z \\
\text{subject to} \quad & \boldsymbol{y} \le \boldsymbol{\alpha} z, \ \boldsymbol{y} \in \mathbb{R}_+^k, \ 0 \le z \le x.
\end{aligned}
$$

As before, we restrict ourselves in the following to the nontrivial case where $\boldsymbol{\alpha}^\top \mathbf{e} \ge \beta$.

Theorem 2. *Determining an ϵ-optimal decision to the linear two-stage stochastic program (9) with fixed recourse is #P-hard whenever $\epsilon < \delta/(8 \max\{1, \boldsymbol{\alpha}^\top \mathbf{e} - \beta\})$ for δ defined in Theorem 1.*

Proof. Note that $\mathbb{E}\left[Q(x, \tilde{\boldsymbol{\xi}}; \boldsymbol{\alpha}, \beta)\right] = \mathcal{Q}(\boldsymbol{\alpha}, \beta)x$, where $\mathcal{Q}(\boldsymbol{\alpha}, \beta)$ denotes the expected recourse value of the second-stage problem (2). Thus, the optimal solution $x^\star$ and the optimal value $f^\star$ of problem (9) satisfy $x^\star = 0$ and $f^\star = 0$ if $c > \mathcal{Q}(\boldsymbol{\alpha}, \beta)$, $x^\star \in [0,1]$ and $f^\star = 0$ if $c = \mathcal{Q}(\boldsymbol{\alpha}, \beta)$, and $x^\star = 1$ and $f^\star = \mathcal{Q}(\boldsymbol{\alpha}, \beta) - c$ if $c < \mathcal{Q}(\boldsymbol{\alpha}, \beta)$. Assume now that an ϵ-optimal decision x to problem (9) could be found in polynomial time, and consider the following bisection search:

1. Set $[\underline{c}, \overline{c}] = [0, \boldsymbol{\alpha}^\top \mathbf{e} - \beta]$.

2. Find the ϵ-optimal solution x to problem (9) with $c = (\underline{c} + \overline{c})/2$.

3. Set $\overline{c} = c$ if $x < \frac{1}{2}$ and set $\underline{c} = c$ otherwise.

4. Go back to Step 2 as long as $\overline{c} - \underline{c} > \delta$.

We claim that $\underline{c} - \frac{\delta}{2} \le \mathcal{Q}(\boldsymbol{\alpha}, \beta) \le \overline{c} + \frac{\delta}{2}$ in every iteration of this algorithm. This is certainly the case in the first iteration. Assume now that the condition is satisfied in the i-th iteration. We then distinguish the two cases $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c| > \frac{\delta}{2}$ and $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c| \le \frac{\delta}{2}$.

Assume first that $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c| > \frac{\delta}{2}$, and let x be an ϵ-optimal solution to problem (9). Since $c, \mathcal{Q}(\boldsymbol{\alpha}, \beta) \in [0, \boldsymbol{\alpha}^\top \mathbf{e} - \beta]$, we have $\max\{1, |\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c|x^\star\} \le \max\{1, \boldsymbol{\alpha}^\top \mathbf{e} - \beta\}$ and thus

$$
\frac{|(\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c)(x^\star - x)|}{\max\{1, |\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c|x^\star\}} \le \epsilon \implies |\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c| \, |x^\star - x| \le \epsilon \max\{1, \boldsymbol{\alpha}^\top \mathbf{e} - \beta\}
$$

$$
\iff |x^\star - x| \le \frac{\epsilon \max\{1, \boldsymbol{\alpha}^\top \mathbf{e} - \beta\}}{|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c|}
$$

$$
\implies |x^\star - x| \le \frac{1}{4},
$$

where the last implication follows from the induction hypothesis and the definition of ϵ. Thus, $x < \frac{1}{2}$ implies that $x^\star = 0$, that is, $c > \mathcal{Q}(\boldsymbol{\alpha}, \beta)$, whereas $x \ge \frac{1}{2}$ implies that $x^\star = 1$ and therefore

9

$c < \mathcal{Q}(\boldsymbol{\alpha}, \beta)$. In either case, we have $\underline{c} - \frac{\delta}{2} \leq \mathcal{Q}(\boldsymbol{\alpha}, \beta) \leq \overline{c} + \frac{\delta}{2}$ in the $(i+1)$-th iteration. Note that the case $c = \mathcal{Q}(\boldsymbol{\alpha}, \beta)$ is excluded since we assumed that $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c| > \frac{\delta}{2}$.

Assume now that $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - c| \leq \frac{\delta}{2}$. If we set $\overline{c} = c$ in the i-th iteration, then $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - \overline{c}| \leq \frac{\delta}{2}$ in the $(i+1)$-th iteration, that is, $\mathcal{Q}(\boldsymbol{\alpha}, \beta) \leq \overline{c} + \frac{\delta}{2}$, while $\underline{c}$ remains unchanged. Likewise, if we set $\underline{c} = c$, then $|\mathcal{Q}(\boldsymbol{\alpha}, \beta) - \underline{c}| \leq \frac{\delta}{2}$ in the next iteration, that is, $\underline{c} - \frac{\delta}{2} \leq \mathcal{Q}(\boldsymbol{\alpha}, \beta)$, while $\overline{c}$ remains unchanged. Thus, the condition $\underline{c} - \frac{\delta}{2} \leq \mathcal{Q}(\boldsymbol{\alpha}, \beta) \leq \overline{c} + \frac{\delta}{2}$ is preserved in both cases. $\qquad\square$

Another consequence of Theorem 1 is that computing the expected value of the non-negative part of a linear combination of uniformly distributed random variables is #P-hard.

Corollary 1. *For fixed values $\boldsymbol{\alpha} \in \mathbb{R}_+^k$ and $\beta \in \mathbb{R}_+$, approximating the expectation*

$$\mathbb{E}\left[\max\left\{\boldsymbol{\alpha}^\top \tilde{\boldsymbol{\xi}} - \beta, 0\right\}\right]$$

to within an absolute accuracy δ that satisfies the condition of Theorem 1 is #P-hard even if $\tilde{\boldsymbol{\xi}}$ follows the uniform distribution supported on $[0,1]^k$.

3   Stochastic Programs with Random Recourse

We now show that approximately computing the optimal value of the two-stage stochastic program (1) with *random* recourse is *strongly* #P-hard, that is, the problem remains #P-hard under a unary encoding [6]. Our proof relies on the auxiliary result that calculating the volume of an order polytope to within high accuracy is #P-hard. For a set $S = \{1, \ldots, k\}$ and a partial order $\preceq$ defined on S, the order polytope is $P(S; \preceq) = \{\boldsymbol{\xi} \in [0,1]^k : \xi_i \leq \xi_j \ \forall i, j \in S, i \preceq j\}$.

Lemma 3. *Computing the volume of the order polytope $P(S; \preceq)$ to within an absolute accuracy of ϵ is #P-hard whenever $\epsilon < \frac{1}{2k!}$.*

The proof of Lemma 3 constructs a reduction from the #LINEAREXTENSION problem, whose #P-hardness has been established in [3].

#LINEAREXTENSION

Instance. Given are a set $S = \{1, \ldots, k\}$ and a partial order $\preceq$ defined on S.

Question. Compute the number of linear extensions $N(S; \preceq)$ of S under $\preceq$.

We remind the reader that a linear extension of S under $\preceq$ is a permutation $i_1, \ldots, i_k$ of the elements in S that satisfies $\ell \leq \ell'$ whenever $i_\ell \preceq i_{\ell'}$.

Proof of Lemma 3. Fix an instance of the #LINEAREXTENSION problem with input S and $\preceq$, and let ϑ be an approximation of $\mathrm{Vol}(P(S; \preceq))$ to within an absolute accuracy of $\epsilon \in [0, 1/2k!)$. It is shown in [3] that the number of linear extensions $N(S; \preceq)$ of S under $\preceq$ satisfies $N(S; \preceq) = k!\,\mathrm{Vol}(P(S; \preceq))$. Thus, by multiplying ϑ with $k!$ and rounding the result to the nearest integer, we can compute $N(S; \preceq)$. Since each step except for the volume calculation can be carried out in time polynomial in the bit length of $(S, \preceq)$, we conclude that approximating the volume of an order polytope is at least as hard as solving the #P-hard #LINEAREXTENSION problem. $\qquad\square$

Lemma 3 allows us to establish the first main result of this section. The result studies the expected recourse value $\mathcal{Q}(S, \preceq) = \mathbb{E}[Q(\tilde{\boldsymbol{\xi}}; S, \preceq)]$ of the second-stage problem

$$
\begin{aligned}
\text{maximize} \quad & z \\
\text{subject to} \quad & z \leq \sum_{i,j \in S : i \preceq j} (\xi_i - \xi_j) y_{ij} \\
& \boldsymbol{y} \in \mathbb{R}_+^{k \times k}, \quad z \in [0, 1]
\end{aligned}
\tag{10}
$$

under the uniform distribution supported on $[0, 1]^k$, where S and $\preceq$ are fixed parameters. Note that problem (10) exhibits random recourse but does not involve any first-stage decisions.

Theorem 3. *Computing the expected recourse value $\mathcal{Q}(S, \preceq)$ of the second-stage problem* (10) *for S and $\preceq$ to within an absolute accuracy of ϵ is strongly #P-hard whenever $\epsilon < \frac{1}{2k!}$.*

Proof. One readily verifies that for a fixed realization $\boldsymbol{\xi}$ of the random vector $\tilde{\boldsymbol{\xi}}$, the optimal value of the second-stage problem (10) is 0 if $\boldsymbol{\xi} \in P(S; \preceq)$ and 1 otherwise. We thus conclude that

$$
\mathcal{Q}(S, \preceq) \;=\; \mathbb{E}[Q(\tilde{\boldsymbol{\xi}}; S, \preceq)] \;=\; \mathbb{E}\left[\mathbb{I}_{[\tilde{\boldsymbol{\xi}} \notin P(S, \preceq)]} \right] \;=\; 1 - \mathrm{Vol}(P(S; \preceq)),
$$

that is, computing the expected recourse value of (10) to within an absolute accuracy of ϵ allows us to compute the volume of the associated order polytope to within the same accuracy. From Lemma 3 we know, however, that the latter problem is #P-hard.

It remains to show that computing the expected recourse value $\mathcal{Q}(S, \preceq)$ is *strongly* #P-hard. Assume, to the contrary, that there is a pseudo-polynomial time algorithm for computing $\mathcal{Q}(S, \preceq)$ to within an accuracy of ϵ, that is, an algorithm whose runtime can be bounded by a polynomial

in the bit length of the description of the affine mappings $\boldsymbol{q}$, $\boldsymbol{T}$, $\boldsymbol{W}$ and $\boldsymbol{h}$ in (1) that correspond to problem (10), as well as the largest numeric value occuring in that description. Since the affine mappings in problem (10) only contain the coefficients -1, 0 and 1, the runtime of such an algorithm would indeed be polynomial in the bit length of the description of problem (10). Thus, the existence of a pseudo-polynomial time algorithm for computing $\mathcal{Q}(S, \preceq)$ to within an accuracy of ϵ would imply that the volume of the associated order polytope could be computed efficiently, which establishes the strong #P-hardness of the problem. $\qquad\square$

Since the expected recourse value of problem (10) can be computed through a linear two-stage stochastic program (1) with random recourse, we conclude that approximating the optimal value of a linear two-stage stochastic program with random recourse to high accuracy is #P-hard as well.

We conclude this section by showing that determining an approximately optimal decision to problem (1) with random recourse is strongly NP-hard. This implies that unless the problems in NP admit an efficient solution scheme, there is no fully polynomial-time approximation scheme (FPTAS) for computing the maximum of problem (1) with random recourse, see [6].

To this end, we recall the definition of the strongly NP-hard Integer Feasibility Problem [6]:

INTEGER FEASIBILITY PROBLEM

Instance. Given are $\boldsymbol{A} \in \mathbb{Z}^{m \times n}$ and $\boldsymbol{b} \in \mathbb{Z}^m$ such that $\{\boldsymbol{y} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{b}\} \subseteq [0, 1]^n$.

Question. Is there a vector $\boldsymbol{y} \in \{0, 1\}^n$ such that $\boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{b}$?

The following lemma shows that we can round fractional solutions to the Integer Feasibility Problem if they are sufficiently close to a binary vector.

Lemma 4. *Fix any $\epsilon' < \min_i \{(\sum_j |A_{ij}|)^{-1}\}$ that satisfies $\epsilon' < \frac{1}{2}$. The Integer Feasibility Problem has an affirmative answer if and only if there is a vector $\boldsymbol{y} \in ([0, \epsilon'] \cup [1 - \epsilon', 1])^n$ such that $\boldsymbol{A}\boldsymbol{y} \leq \boldsymbol{b}$.*

Proof. The 'only if' direction is immediate. As for the 'if' direction, fix a fractional vector $\boldsymbol{y}$ as described in the statement and let $\boldsymbol{y}'$ be the closest binary vector, that is, $y_i' := 1$ if $y_i \geq 1 - \epsilon'$; $:= 0$ if $y_i \leq \epsilon'$. We then observe that $\sum_j A_{ij} y_j' \leq \sum_j A_{ij} y_j + \sum_j |A_{ij}| \epsilon' < \sum_j A_{ij} y_j + 1 \leq b_i + 1$ for all $i = 1, \ldots, m$. Due to the integrality of $\boldsymbol{A}$, $\boldsymbol{y}'$ and $\boldsymbol{b}$, we thus conclude that $\boldsymbol{A}\boldsymbol{y}' \leq \boldsymbol{b}$. $\qquad\square$

In the following, we fix an instance $(\boldsymbol{A}, \boldsymbol{b})$ of the Integer Feasibility Problem and consider the

following instance of problem (1) with random recourse:

$$
\begin{aligned}
\text{minimize} \quad & x + \mathbb{E}\left[Q(x, \tilde{\boldsymbol{\xi}})\right] \\
\text{subject to} \quad & x \in \mathbb{R},
\end{aligned}
\tag{11}
$$

where $\tilde{\boldsymbol{\xi}}$ follows a uniform distribution supported on $[0, 1]^k$, and for a fixed realization $\boldsymbol{\xi}$ the recourse function $Q(x, \boldsymbol{\xi})$ amounts to the optimal value of the second-stage problem

$$
\begin{aligned}
\text{minimize} \quad & \mathbf{e}^\top \boldsymbol{y} \\
\text{subject to} \quad & \boldsymbol{y} \in \mathbb{R}^n_+, \ \boldsymbol{\lambda} \in \mathbb{R}^m_+, \ x \geq \mathbf{e}^\top \boldsymbol{y} \\
& \left. \begin{aligned} y_i &\geq \xi_i + (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{\xi})^\top \boldsymbol{\lambda} \\ y_i &\geq (1 - \xi_i) + (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{\xi})^\top \boldsymbol{\lambda} \end{aligned} \right\} \quad \forall i = 1, \ldots, m.
\end{aligned}
$$

Note that this problem does not have relatively complete recourse, that is, for small values of x the second-stage problem may become infeasible with positive probability, which in turn implies that this choice of x is infeasible in the first-stage problem (11). Problems without relatively complete recourse are common in stochastic programming, see [2]. Also, the problem satisfies $n = m = k$.

Theorem 4. *Determining an ϵ-optimal decision to the linear two-stage stochastic program* (11) *with random recourse is strongly* NP*-hard whenever $\epsilon < \epsilon'/4n$ for ϵ' defined in Lemma 4.*

Proof. If the second-stage problem is feasible for a fixed $\boldsymbol{\xi}$, then any optimal solution $(\boldsymbol{y}(\boldsymbol{\xi}), \boldsymbol{\lambda}(\boldsymbol{\xi}))$ to the second-stage problem satisfies $y_i(\boldsymbol{\xi}) = \max\{\xi_i, 1 - \xi_i\}$, $i = 1, \ldots, n$, if $\boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b}$ and $\boldsymbol{y}(\boldsymbol{\xi}) = \boldsymbol{0}$ otherwise. Assuming that $\{\boldsymbol{\xi} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b}\} \neq \emptyset$, the optimal solution $x^\star$ to problem (11) thus satisfies $x^\star = \max\{\sum_{i=1}^m \max\{\xi_i, 1 - \xi_i\} : \boldsymbol{A}\boldsymbol{\xi} \leq \boldsymbol{b}\}$. Lemma 4 then implies that the answer to the Integer Feasibility Problem is affirmative if and only if $x^\star = n$, and it is negative if and only if $x^\star < n - \epsilon'$.

Assume now that an ϵ-optimal solution x to problem (11) could be found in polynomial time. In that case, we obtain

$$
\begin{aligned}
& \frac{\left| x^\star + \mathbb{E}\left[Q(x^\star, \tilde{\boldsymbol{\xi}})\right] - x - \mathbb{E}\left[Q(x, \tilde{\boldsymbol{\xi}})\right] \right|}{\max\left\{1, x^\star + \mathbb{E}\left[Q(x^\star, \tilde{\boldsymbol{\xi}})\right]\right\}} \ \leq \ \epsilon \\
\iff \quad & |x^\star - x| \ \leq \ \epsilon \cdot \max\left\{1, x^\star + \mathbb{E}\left[Q(x^\star, \tilde{\boldsymbol{\xi}})\right]\right\} \\
\implies \quad & |x^\star - x| \ \leq \ 2n\epsilon \ < \ \frac{\epsilon'}{2},
\end{aligned}
$$

where the equivalence follows from the fact that $\mathbb{E}\left[Q(x^\star, \tilde{\boldsymbol{\xi}})\right] = \mathbb{E}\left[Q(x, \tilde{\boldsymbol{\xi}})\right]$ since $x^\star$ and x are both feasible, the first implication holds since both $x^\star$ and $\mathbb{E}\left[Q(x^\star, \tilde{\boldsymbol{\xi}})\right]$ do not exceed n by construction, and the last implication follows from the definition of ϵ. We thus conclude that $x^\star = n$ whenever $x > n - \frac{\epsilon'}{2}$ and $x^\star < n - \epsilon'$ whenever $x < n - \frac{\epsilon'}{2}$, that is, we could decide the NP-hard Integer Feasibility Problem in polynomial time. $\qquad\square$

4   Conclusion

Despite the hardness results addressed in this paper, there are many interesting two-stage stochastic programs that can be solved to within workable accuracy. In [12] it is argued, for instance, that two-stage stochastic programs with relatively complete recourse can be solved efficiently with the sample average approximation method to a moderate relative accuracy of 10% or even 1%. Recall that a two-stage stochastic program has relatively complete recourse if for every feasible first-stage decision and for every possible realization of the uncertain parameters there exists a feasible second-stage decision. We emphasize, however, that problem (1) was not assumed to have relatively complete recourse. Modern methods of Quasi-Monte Carlo (QMC) integration provide other effective tools for solving two-stage stochastic programs. Specifically, it is known that integration via QMC algorithms overcomes the curse of dimensionality in certain mixed Sobolev spaces [13]. This means that the number of function evaluations needed to guarantee a relative accuracy of ϵ across all integrands in the underlying Sobolev space is at most polynomial in ϵ^{-1} and the dimension of $\boldsymbol{\xi}$. This result is consistent with the complexity theorems discussed in this note because the optimal value functions $Q(\boldsymbol{x}, \boldsymbol{\xi})$ of linear two-stage stochastic programs fail to belong to the relevant Sobolev spaces. Moreover, the critical accuracies given in Theorems 1 and 3, beyond which the evaluation of the expected value functions becomes intractable, are exponentially small in the input sizes of the corresponding stochastic programs. Under relatively complete recourse and other technical conditions, it can be shown, however, that $Q(\boldsymbol{x}, \boldsymbol{\xi})$ can be approximated by appropriate Sobolev integrands, which in turn implies that $Q(\boldsymbol{x}, \boldsymbol{\xi})$ can be integrated efficiently via QMC algorithms with attractive convergence guarantees independent of the dimension of $\boldsymbol{\xi}$ [8].

Acknowledgments

References

[1] N. Balakrishnan and V. B. Nevzorov. *A Primer on Statistical Distributions*. John Wiley & Sons, 2004.

[2] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer Series in Operations Research. Springer, 1997.

[3] G. Brightwell and P. Winkler. Counting linear extensions. *Order*, 8(3):225–242, 1991.

[4] M. Dyer and L. Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming A*, 106(3):423–432, 2006.

[5] M. E. Dyer and A. M. Frieze. On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing*, 17(5):967–974, 1988.

[6] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[7] W. Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. *Numerische Mathematik*, 4(1):117–123, 1962.

[8] H. Leövey and W. Römisch. Quasi-Monte Carlo methods for linear two-stage stochastic programming problems. *Mathematical Programming*, 151(1):315–345, 2015.

[9] C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.

[10] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, 1995.

[11] A. Ruszczyński and A. Shapiro, editors. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*. Elsevier, 2003.

[12] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. In V. Jeyakumar and A.M. Rubinov, editors, *Continuous Optimization: Current Trends and Applications*, pages 111–144. Springer, 2005.

[13] I. H. Sloan and H. Woźniakowski. When are Quasi-Monte Carlo algorithms efficient for high dimensional integrals? *Journal of Complexity*, 14(1):1–33, 1998.