

# Volumetric barrier decomposition algorithms for two-stage stochastic linear semi-infinite programming

Baha Alzalg<sup>1,2</sup> · Asma Gafour<sup>1,3</sup> · Lewa' Alzaleq<sup>4</sup> \*

December 6, 2018

**Abstract** In this paper, we study the two-stage stochastic linear semi-infinite programming with recourse to handle uncertainty in data defining (deterministic) linear semi-infinite programming. We develop and analyze volumetric barrier decomposition-based interior point methods for solving this class of optimization problems, and present a complexity analysis of the proposed algorithms. We establish our convergence analysis by showing that the volumetric barrier associated with the recourse function of stochastic linear semi-infinite programs is a strongly self-concordant barrier and forms a self-concordant family on the first-stage solutions. The dominant terms in the complexity expressions obtained in this paper are given in terms of the problem dimension and the number of realizations. The novelty of our algorithms lies in their ability to kill the effect of the radii of the largest Euclidean balls contained in the feasibility sets on the dominant complexity terms.

**Keywords:** Semi-infinite programming, Stochastic linear programming, Interior point methods, Volumetric barrier, Decomposition

**AMS Classification:** 90C06, 90C15, 90C34, 90C51, 90C60

## 1 Introduction

The purpose of this paper is to introduce and study the two-stage stochastic linear semi-infinite programming (SLSIP in brief) with recourse in the dual standard form

$$\begin{aligned} \max \quad & c^\top x + \mathbb{E} [Q(x, \omega)] \\ \text{s.t.} \quad & a_\gamma^\top x \leq b_\gamma, \quad \forall \gamma \in \Gamma, \end{aligned} \tag{1}$$

---

✉ Baha Alzalg (Corresponding author)  
b.alzalg@ju.edu.jo, bmnsma@rit.edu

Asma Gafour  
asma\_gafour@yahoo.fr

Lewa Alzaleq  
lewa.alzaleq@wsu.edu

1 Department of Mathematics, The University of Jordan, Amman, 11942, Jordan

2 School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY, 14623, USA

3 Department of Mathematics, University of Djilali Liabes, Sidi Bel Abbes, 22001, Algeria

4 Department of Mathematics and Statistics, Washington State University, Pullman, WA, 99164, USA

where  $x \in \mathbb{R}^{m_1}$  is the first-stage decision variable,  $\Gamma$  is an index set, the vectors  $c, a_\gamma \in \mathbb{R}^{m_1}$  and  $b_\gamma \in \mathbb{R}$  ( $\gamma \in \Gamma$ ) are deterministic data, and  $Q(x, \omega)$  is the maximum value of the problem

$$\begin{aligned} \max \quad & d(\omega)^\top \mathbf{y} \\ \text{s.t.} \quad & \mathbf{w}_\lambda(\omega)^\top \mathbf{y} \leq h_\lambda(\omega) - \mathbf{t}_\lambda(\omega)^\top \mathbf{x}, \quad \forall \lambda \in \Lambda, \end{aligned} \quad (2)$$

where  $\mathbf{y} \in \mathbb{R}^{m_2}$  is the second-stage variable,  $\Lambda$  is an index set, the vectors  $d(\omega), \mathbf{w}_\lambda(\omega), \mathbf{t}_\lambda(\omega) \in \mathbb{R}^{m_2}$  and  $h_\lambda(\omega) \in \mathbb{R}$  ( $\lambda \in \Lambda$ ) are random data whose realizations depend on an underlying outcome  $\omega$  in an event space  $\Omega$  with a known probability function  $\mathcal{P}$ , and

$$\mathbb{E}[Q(x, \omega)] := \int_{\Omega} Q(x, \omega) \mathcal{P}(d\omega).$$

To the best of our knowledge, the optimization problem introduced in (1) and (2) has not been studied yet, although it is the simplest nontrivial stochastic semi-infinite optimization problem. A linear semi-infinite program is an optimization problem with a linear objective function and linear constraints in which either the number of unknowns or the number of constraints is finite. Clearly, the SLSIP problem (1) and (2) has a finite number of unknowns and an infinite number of constraints.

It can be seen that the SLSIP generalizes the ordinary stochastic linear programming by allowing infinite number of constraints on one hand, and the deterministic linear semi-infinite programming by allowing uncertainty in data on the other hand. The SLSIP is also a special case of the stochastic semi-infinite programming (or, more generally, the stochastic infinite-dimensional programming) by enforcing linearity in the objective function and the constraints. The very broad and direct applications of each of linear programs, stochastic programs [1,2] and semi-infinite programs [3,4] attracted us to study and analyze SLSIPs as a promising class of optimization problems applicable to a wide range of real-life problems. See the models in [5–7] which describe some applications in radiation transfer theory, neutron transport theory, and waste management.

Luo et al. [8] (see also [9,10]) derived a logarithmic barrier decomposition-based interior point algorithm for deterministic linear semi-infinite programming. An alternative to the logarithmic barrier is the volumetric barrier of Vaidya [11] (see also [12]). It has been found [13] that some cutting plane algorithms for stochastic linear programming problems based on the volumetric barrier perform practically better than those based on the logarithmic barrier. For this reason, a number of volumetric barrier Benders' decomposition-based interior point algorithms have been developed recently for solving stochastic (linear, second-order, and semidefinite) cone optimization problems. Below we briefly outline these algorithmic results. In 2007, Ariyawansa and Zhu [14] derived a volumetric barrier decomposition interior point algorithm for two-stage stochastic (convex) quadratic linear programming. In 2011, Ariyawansa and Zhu [15] generalized their work in [14] to derive a volumetric barrier decomposition interior point algorithm for two-stage stochastic semidefinite programming. In 2015, Alzalg [16] exploited the work of Ariyawansa and Zhu in [14,15] to derive a volumetric barrier decomposition interior point algorithm for two-stage stochastic second-order cone programming.

Note that the setting in this paper is similar to that of Ariyawansa and Zhu [14] for stochastic quadratic linear programs, but the linearity is assumed in the objective function, and, most notably, the semi-infiniteness is involved in the constraints of our setting. The current setting is also similar to that of Luo et al. [8] for deterministic linear semi-infinite programming but the stochasticity is assumed here. In this paper, we utilize the work of Luo et al. [8] for deterministic linear semi-infinite programming on one hand and the work of Ariyawansa and Zhu [14] for ordinary stochastic quadratic linear programming on the other hand to derive volumetric barrier decomposition-based interior point algorithms for two-stage SLSIP problem with recourse.

We establish our convergence analysis by showing that the volumetric barrier associated with the recourse function of stochastic linear semi-infinite programs behaves as a strongly self-concordant barrier and forms a self-concordant family on the first-stage solutions. We will see that the self-concordance analysis is established in this work in a different way than that in Section 3 of [14]. In fact, as (convex) quadratic linear programming is a special case of semidefinite programming, the authors in [14] re-wrote the stochastic quadratic linear programming problem as a stochastic semidefinite programming problem (by formulating the linear inequalities as linear matrix inequalities) and heavily made use of their own results in [15, Section 4] for stochastic semidefinite programming. In comparison, we approach the self-concordance proofs from a linear programming point of view, which gives more explicit and direct proofs because we use more elementary arguments.

We establish polynomial complexity of the resulting methods. The dominant terms of the complexity expressions obtained in this work are given in terms of the problem dimension and the number of realizations. Unlike the complexity expression obtained for the logarithmic barrier algorithm in [8], the dominant complexity terms for our volumetric barrier algorithm are not affected by the radii of the largest Euclidean balls contained in the feasibility sets. From this advantage comes the importance of the development of this paper. We will see that this significant advantage stems from the use of the volumetric barrier instead of using the logarithmic barrier. We will also see that the “rich flavor” hidden inside the volumetric barrier can be tasted in the proposed algorithm of this work more than in their counterparts algorithms in [14–16].

We mention how this paper is structured. In Section 2, we present our problem formulation, discretization, some assumptions, and introduce the volumetric barrier problem associated with the SLSIP problem. In Section 3, we first show that the problem of finding an approximate minimizer of the SLSIP problem can be reduced to that of finding an approximate minimizer of the discretized problem under a certain condition on the measure of proximity of the current point  $x$  to the central path. Then, we compute the gradient and Hessian of the barrier functions in the second part of Section 3. In Section 4, we show that the set of volumetric barrier functions for positive values of barrier parameter forms a self-concordant family. Based on this property, we present a class of volumetric barrier decomposition-based interior point algorithms and provide their convergence and complexity in Section 5. Section 6 contains some concluding remarks. The proofs of the convergence and complexity results are given in the Appendix A.

We end this section by introducing some notations that will be used in the sequel. Let  $\mathbb{R}^{m \times n}$  denote the vector space of real  $m \times n$  matrices. We use  $\odot$  to denote the Hadamard product of matrices; i.e.  $(U \odot V)_{ij} = u_{ij}v_{ij}$  for  $U, V \in \mathbb{R}^{m \times n}$ . Let  $\mathbb{R}^{n \times n}$  denote the vector space of real symmetric  $n \times n$  matrices. For  $U, V \in \mathbb{R}^{n \times n}$ , we write  $U \geq 0$  ( $U > 0$ ) to mean that  $U$  is positive semidefinite (positive definite) and  $U \geq V$  or  $V \leq U$  to mean that  $U - V \geq 0$ . For any strictly positive vector  $x \in \mathbb{R}^n$ , we define  $\ln x := (\ln x_1, \dots, \ln x_n)^\top$ ,  $\sqrt{x} := (\sqrt{x_1}, \dots, \sqrt{x_n})^\top$  and  $x^{-1} := (x_1^{-1}, \dots, x_n^{-1})^\top$ . We also use  $X := \text{diag}(x_1, \dots, x_n)$  to denote the  $n \times n$  diagonal matrix whose diagonal entries are  $x_1, \dots, x_n$ .

## 2 Problem formulation and assumptions

In this section, we first present the extensive formulation of the SLSIP problem (1) and (2) with discretization. Then, we present the volumetric barrier problem for SLSIPs.

## 2.1 The SLSIP problem formulation and discretization

We examine (1) and (2) when the event space  $\Omega$  is finite. Let

$$\{(\mathbf{d}_0^{(k)}, \mathbf{w}_\lambda^{(k)}, \mathbf{t}_\lambda^{(k)}, h_\lambda^{(k)}) : k = 1, \dots, K, \lambda \in \Lambda\}$$

be the set of the possible values of the random variables  $(\mathbf{d}_0(\omega), \mathbf{w}_\lambda(\omega), \mathbf{t}_\lambda(\omega), h_\lambda(\omega))$  and let

$$p_k := \mathcal{P}(\mathbf{d}_0(\omega), \mathbf{w}_\lambda(\omega), \mathbf{t}_\lambda(\omega), h_\lambda(\omega)) = (\mathbf{d}_0^{(k)}, \mathbf{w}_\lambda^{(k)}, \mathbf{t}_\lambda^{(k)}, h_\lambda^{(k)})$$

be the associated probability for  $k = 1, 2, \dots, K$ . Then, the two-stage SLSIP problem (1) and (2) becomes

$$\begin{aligned} \max \quad & \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^K p_k \varrho^{(k)}(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{a}_\gamma^\top \mathbf{x} \leq b_\gamma, \quad \forall \gamma \in \Gamma, \end{aligned} \quad (3)$$

where, for  $k = 1, 2, \dots, K$ ,  $\varrho^{(k)}(\mathbf{x})$  is the maximum value of the problem

$$\begin{aligned} \max \quad & \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} \\ \text{s.t.} \quad & \mathbf{w}_\lambda^{(k)\top} \mathbf{y}^{(k)} \leq h_\lambda^{(k)} - \mathbf{t}_\lambda^{(k)\top} \mathbf{x}, \quad \forall \lambda \in \Lambda, \end{aligned} \quad (4)$$

We re-define  $\mathbf{d}^{(k)}$  as  $\mathbf{d}^{(k)} := p_k \mathbf{d}^{(k)}$  for  $k = 1, 2, \dots, K$ . Further, in order to conveniently make use of the results in [12], we rewrite the SLSIP problem (3) and (4) in the following equivalent form:

$$\begin{aligned} \min \quad & \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^K \varrho^{(k)}(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{a}_\gamma^\top \mathbf{x} \geq b_\gamma, \quad \forall \gamma \in \Gamma, \end{aligned} \quad (5)$$

where, for  $k = 1, 2, \dots, K$ ,  $\varrho^{(k)}(\mathbf{x})$  is the minimum value of the problem

$$\begin{aligned} \min \quad & \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} \\ \text{s.t.} \quad & \mathbf{w}_\lambda^{(k)\top} \mathbf{y}^{(k)} \geq h_\lambda^{(k)} - \mathbf{t}_\lambda^{(k)\top} \mathbf{x}, \quad \forall \lambda \in \Lambda. \end{aligned} \quad (6)$$

Without loss of generality, we assume the vectors  $\mathbf{c}$  and  $\mathbf{a}_\gamma, \gamma \in \Gamma$ , are normalized so that  $\|\mathbf{c}\| = \|\mathbf{a}_\gamma\| = 1$ , for all  $\gamma \in \Gamma$ . We also assume the vectors  $\mathbf{d}^{(k)}$  and  $\mathbf{w}_\lambda^{(k)}$  and  $\mathbf{t}_\lambda^{(k)}, \gamma \in \Gamma$ , are normalized so that  $\|\mathbf{d}^{(k)}\| = \|\mathbf{w}_\lambda^{(k)}\| = \|\mathbf{t}_\lambda^{(k)}\| = 1$ , for all  $\lambda \in \Lambda$  and  $k = 1, 2, \dots, K$ . Let

$$\begin{aligned} \mathcal{F}_1 &:= \{\mathbf{x} : \mathbf{a}_\gamma^\top \mathbf{x} \geq b_\gamma, \gamma \in \Gamma\}; \\ \mathcal{F}_2^{(k)}(\mathbf{x}) &:= \{\mathbf{y}^{(k)} : \mathbf{w}_\lambda^{(k)\top} \mathbf{y}^{(k)} \geq h_\lambda^{(k)} - \mathbf{t}_\lambda^{(k)\top} \mathbf{x}\} \text{ for } k = 1, 2, \dots, K; \\ \mathcal{F}_2 &:= \{\mathbf{x} : \mathcal{F}_2^{(k)}(\mathbf{x}) \neq \emptyset, k = 1, 2, \dots, K\}; \\ \mathcal{F} &:= \mathcal{F}_1 \cap \mathcal{F}_2. \end{aligned}$$

Then  $\mathcal{F}$  denotes the feasible set of the SLSIP problem (5) and (6). Throughout the paper, we make the assumptions that  $\mathcal{F}_1$  is contained in the unit hyperbolic  $[0, 1]^{m_1}$ ,  $\mathcal{F}_2^{(k)}(\mathbf{x})$  is contained in the unit hyperbolic  $[0, 1]^{m_2}$  and  $\mathcal{F}$  has nonempty interior.

For any subsets  $Q_1 \subset \Gamma$  and  $Q_2 \subset \Lambda$ , we can define a corresponding discretization (or relaxation) of the SLSIP problem (5) and (6) by considering only those constraints indexed by  $Q_1$  in the

subproblem (5) and only those constraints indexed by  $Q_2$  in the subproblem (6). So, we can consider the following discretization of the SLSIP problem (5) and (6).

$$\begin{aligned} \min \quad & c^\top x + \sum_{k=1}^K \varrho^{(k,n_2)}(x) \\ \text{s.t.} \quad & a_i^\top x \geq b_i, \quad i = 1, 2, \dots, n_1, \end{aligned} \quad (7)$$

where, for  $k = 1, 2, \dots, K$ ,  $\varrho^{(k,n_2)}(x)$  is the minimum value of the problem

$$\begin{aligned} \min \quad & d^{(k)\top} y^{(k)} \\ \text{s.t.} \quad & w_j^{(k)\top} y^{(k)} \geq h_j^{(k)} - t_j^{(k)\top} x, \quad j = 1, 2, \dots, n_2. \end{aligned} \quad (8)$$

We construct the above discretization by choosing  $n_1$  ( $n_1 \geq 2m_1$ ) linear constraints from the constraint set  $\{a_\gamma^\top x \geq b_\gamma : \gamma \in \Gamma\}$  and  $n_2$  ( $n_2 \geq 2m_2$ ) linear constraints from the constraint set  $\{w_\lambda^{(k)\top} y^{(k)} \geq h_\lambda^{(k)} - t_\lambda^{(k)\top} x : \lambda \in \Lambda\}$  for each  $k = 1, 2, \dots, K$ . As will be illustrated thoroughly later in the next section, the problem of finding an approximate minimizer of the SLSIP problem (5) and (6) can be reduced to that of finding an approximate solution:  $x \in \mathcal{F}$ ,  $y^{(1)} \in \mathcal{F}_2^{(1)}(x), \dots, y^{(K)} \in \mathcal{F}_2^{(K)}(x)$  of the discretized problem (7) and (8) provided a certain bound on the measure of proximity of the current point  $x$  to the central path holds.

For the sake of simplicity in presentation, we write  $A^{(n_1)} \in \mathbb{R}^{m_1 \times n_1}$  to denote the matrix whose  $i^{\text{th}}$  column is the vector  $a_i \in \mathbb{R}^{m_1}$  for  $i = 1, 2, \dots, n_1$ . Likewise,  $W^{(k,n_2)} \in \mathbb{R}^{m_2 \times n_2}$  and  $T^{(k,n_2)} \in \mathbb{R}^{m_1 \times n_2}$  denote the matrices whose  $j^{\text{th}}$  columns are the vectors  $w_j^{(k)} \in \mathbb{R}^{m_2}$  and  $t_j^{(k)} \in \mathbb{R}^{m_1}$ , respectively, for  $j = 1, 2, \dots, n_2$  and  $k = 1, 2, \dots, K$ . We also write  $b^{(n_1)} \in \mathbb{R}^{n_1}$  for the vector whose the  $i^{\text{th}}$  component is  $b_i$  for  $i = 1, 2, \dots, n_1$ , and write  $h^{(k,n_2)} \in \mathbb{R}^{n_2}$  for the vector whose the  $j^{\text{th}}$  component is  $h_j^{(k)}$ , for  $j = 1, 2, \dots, n_2$  and  $k = 1, 2, \dots, K$ . With these simplified notations, the SLSIP problem (5) and (6) becomes

$$\begin{aligned} \min \quad & c^\top x + \sum_{k=1}^K \varrho^{(k,n_2)}(x) \\ \text{s.t.} \quad & A^{(n_1)\top} x \geq b^{(n_1)}, \end{aligned} \quad (9)$$

where, for  $k = 1, 2, \dots, K$ ,  $\varrho^{(k,n_2)}(x)$  is the minimum value of the problem

$$\begin{aligned} \min \quad & d^{(k)\top} y^{(k)} \\ \text{s.t.} \quad & W^{(k,n_2)\top} y^{(k)} \geq h^{(k,n_2)} - T^{(k,n_2)\top} x. \end{aligned} \quad (10)$$

The SLSIP problem (9) and (10) can be equivalently written as a DSILP:

$$\begin{aligned} \min \quad & c^\top x + \sum_{k=1}^K d^{(k)\top} y^{(k)} \\ \text{s.t.} \quad & A^{(n_1)\top} x \geq b^{(n_1)} \\ & W^{(k,n_2)\top} y^{(k)} \geq h^{(k,n_2)} - T^{(k,n_2)\top} x, \quad k = 1, 2, \dots, K. \end{aligned} \quad (11)$$

Note that the dual of the SLSIP problem (11) is the problem

$$\begin{aligned}
\max \quad & \mathbf{b}^{(n_1)\top} \boldsymbol{\nu} + \sum_{k=1}^K \left( \mathbf{h}^{(k,n_2)} - T^{(k,n_2)\top} \mathbf{x} \right)^\top \boldsymbol{\nu}^{(k)} \\
\text{s.t.} \quad & A^{(n_1)} \boldsymbol{\nu} + \sum_{k=1}^K T^{(k,n_2)} \boldsymbol{\nu}^{(k)} = \mathbf{c} \\
& W^{(k,n_2)} \boldsymbol{\nu}^{(k)} = \mathbf{d}^{(k)}, \quad k = 1, 2, \dots, K \\
& \boldsymbol{\nu} \geq \mathbf{0}, \quad \boldsymbol{\nu}^{(k)} \geq \mathbf{0}, \quad k = 1, 2, \dots, K,
\end{aligned}$$

where  $\boldsymbol{\nu}$  is the first-stage dual multiplier and  $(\boldsymbol{\nu}^{(1)}; \dots; \boldsymbol{\nu}^{(K)})$  is the second-stage dual multiplier.

## 2.2 The volumetric barrier problem for SLSIPs

First, we define

$$\begin{aligned}
\mathcal{F}_1^\circ &:= \{ \mathbf{x} : \mathbf{s}_1^{(n_1)}(\mathbf{x}) := A^{(n_1)\top} \mathbf{x} - \mathbf{b}^{(n_1)} > \mathbf{0} \}; \\
\mathcal{F}_2^{\circ(k)}(\mathbf{x}) &:= \{ \mathbf{y}^{(k)} : \mathbf{s}_2^{(k,n_2)}(\mathbf{x}, \mathbf{y}^{(k)}) := W^{(k,n_2)\top} \mathbf{y}^{(k)} + T^{(k,n_2)\top} \mathbf{x} - \mathbf{h}^{(k,n_2)} > \mathbf{0} \} \text{ for } k = 1, 2, \dots, K; \\
\mathcal{F}_2^\circ &:= \{ \mathbf{x} : \mathcal{F}_2^{\circ(k)}(\mathbf{x}) \neq \emptyset, k = 1, 2, \dots, K \}; \\
\mathcal{F}^\circ &:= \mathcal{F}_1^\circ \cap \mathcal{F}_2^\circ.
\end{aligned}$$

For simplicity, we write  $\mathbf{s}_1$  and  $\mathbf{s}_2$  for  $\mathbf{s}_1^{(n_1)}(\mathbf{x})$  and  $\mathbf{s}_2^{(k,n_2)}(\mathbf{x}, \mathbf{y}^{(k)})$ , respectively, when it does not lead to confusion. Then we make the following assumptions:

**Assumption 2.1.** The matrices  $A^{(n_1)}$ ,  $T^{(k,n_2)}$  and  $W^{(k,n_2)}$  for all  $k$  have full row rank.

**Assumption 2.2.** The set  $\mathcal{F}^\circ$  is nonempty.

Assumption 2.1 is for convenience. Assumption 2.2 guarantees strong duality for first- and second-stage SLSIPs.

The *logarithmic barrier* for  $\mathcal{F}_1^\circ$  is the function  $\ell_1^{(n_1)} : \mathcal{F}_1^\circ \rightarrow \mathbb{R}$  defined by

$$\ell_1^{(n_1)}(\mathbf{x}) := - \sum_{i=1}^{n_1} \ln s_{1_i}(\mathbf{x}), \quad \forall \mathbf{x} \in \mathcal{F}_1^\circ,$$

and the *volumetric barrier* for  $\mathcal{F}_1^\circ$  is the function  $v_1^{(n_1)} : \mathcal{F}_1^\circ \rightarrow \mathbb{R}$  defined by

$$v_1^{(n_1)}(\mathbf{x}) := \frac{1}{2} \ln \det \left( \nabla_{\mathbf{x}\mathbf{x}}^2 \ell_1^{(n_1)}(\mathbf{x}) \right), \quad \forall \mathbf{x} \in \mathcal{F}_1^\circ.$$

Also under Assumption 2.2,  $\mathcal{F}_2^\circ$  is nonempty and for  $\mathbf{x} \in \mathcal{F}_2^\circ$ ,  $\mathcal{F}_2^{\circ(k)}(\mathbf{x})$  is nonempty for  $k = 1, 2, \dots, K$ . The logarithmic barrier for  $\mathcal{F}_2^{\circ(k)}(\mathbf{x})$  is the function  $\ell_2^{(k,n_2)} : \mathcal{F}_2^{\circ(k)}(\mathbf{x}) \rightarrow \mathbb{R}$  defined by

$$\ell_2^{(k,n_2)}(\mathbf{x}, \mathbf{y}^{(k)}) := - \sum_{j=1}^{n_2} \ln s_{2_j}^{(k)}(\mathbf{x}, \mathbf{y}^{(k)}), \quad \forall \mathbf{y}^{(k)} \in \mathcal{F}_2^{\circ(k)}(\mathbf{x}), \mathbf{x} \in \mathcal{F}_2^\circ,$$

and the volumetric barrier for  $\mathcal{F}_2^{\circ(k)}(\mathbf{x})$  is the function  $v_2^{(k,n_2)} : \mathcal{F}_2^{\circ(k)}(\mathbf{x}) \rightarrow \mathbb{R}$  defined by

$$v_2^{(k,n_2)}(\mathbf{x}, \mathbf{y}^{(k)}) := \frac{1}{2} \ln \det \left( \nabla_{\mathbf{y}^{(k)} \mathbf{y}^{(k)}}^2 \ell_2^{(k,n_2)}(\mathbf{x}, \mathbf{y}^{(k)}) \right), \quad \forall \mathbf{y}^{(k)} \in \mathcal{F}_2^{\circ(k)}(\mathbf{x}), \mathbf{x} \in \mathcal{F}_2^\circ.$$

The volumetric barrier problem associated with the SLSIP problem (9) and (10) is the problem

$$\min \eta^{(n_1, n_2)}(\mu, \mathbf{x}) := \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^K \rho^{(k, n_2)}(\mu, \mathbf{x}) + \mu \varsigma_1 v_1^{(n_1)}(\mathbf{x}), \quad (12)$$

where, for  $k = 1, 2, \dots, K$ ,  $\rho^{(k, n_2)}(\mu, \mathbf{x})$  is the minimum value of the problem

$$\min \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} + \mu \varsigma_2 v_2^{(k, n_2)}(\mathbf{x}, \mathbf{y}^{(k)}). \quad (13)$$

Here  $\varsigma_1 > 0$  and  $\varsigma_2 > 0$  are constants whose values will be defined more precisely in the sequel, and  $\mu > 0$  is a barrier parameter. If for some  $k$ , Problem (13) is infeasible, then we define  $\sum_{k=1}^K \rho^{(k, n_2)}(\mu, \mathbf{x}) := \infty$ .

The SLSIP problem (12) and (13) can be equivalently written as a DSILP:

$$\min \underbrace{\mathbf{c}^\top \mathbf{x} + \mu \varsigma_1 v_1^{(n_1)}(\mathbf{x})}_{\eta^{(n_1, n_2)}(\mu, \mathbf{x})} + \sum_{k=1}^K \underbrace{\left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} + \mu \varsigma_2 v_2^{(k, n_2)}(\mathbf{x}, \mathbf{y}^{(k)}) \right)}_{\rho^{(k, n_2)}(\mu, \mathbf{x})}. \quad (14)$$

Throughout the paper, we denote the optimal solution of Problem (14) by  $(\mathbf{x}(\mu), \mathbf{s}(\mu), \mathbf{y}^{(1)}(\mu, \mathbf{x}), \mathbf{s}_2^{(1)}(\mu, \mathbf{x}), \dots, \mathbf{y}^{(K)}(\mu, \mathbf{x}), \mathbf{s}_2^{(K)}(\mu, \mathbf{x}))$ . The central path is defined as the solution set  $(\mathbf{y}^{(k)}(\mu, \mathbf{x}), \mathbf{s}_2^{(k)}(\mu, \mathbf{x}))$  for  $\mu \geq 0$  and  $\mathbf{x} = \mathbf{x}(\mu)$ .

**Proposition 2.1.** *Let  $\mu > 0$  be fixed. Then  $(\mathbf{x}(\mu), \mathbf{s}(\mu); \mathbf{y}^{(1)}(\mu), \mathbf{s}_2^{(1)}(\mu); \dots; \mathbf{y}^{(K)}(\mu), \mathbf{s}_2^{(K)}(\mu))$  is the optimal solution of (14) if and only if  $(\mathbf{x}(\mu), \mathbf{s}(\mu))$  is the optimal solution for (12) and  $(\mathbf{y}^{(1)}(\mu), \mathbf{s}_2^{(1)}(\mu); \dots; \mathbf{y}^{(K)}(\mu), \mathbf{s}_2^{(K)}(\mu))$  are the optimal solutions for (13) for given  $\mu$  and  $\mathbf{x} = \mathbf{x}(\mu)$ .*

In the next section, we compute the gradient  $\nabla_{\mathbf{x}} \eta(\mu, \mathbf{x})$  and Hessian  $\nabla_{\mathbf{xx}}^2 \eta(\mu, \mathbf{x})$  which are used to determine the *Newton direction* defined by

$$\Delta := \Delta \mathbf{x} := - \left\{ \nabla_{\mathbf{xx}}^2 \eta^{(n_1, n_2)}(\mu, \mathbf{x}) \right\}^{-1} \nabla_{\mathbf{x}} \eta^{(n_1, n_2)}(\mu, \mathbf{x}) \quad (15)$$

for our algorithms. We can also determine the *measure of proximity* of the current point  $\mathbf{x}$  to the central path, which is defined by

$$\delta := \delta(\mu, \mathbf{x}) := \sqrt{\frac{1}{\mu} \Delta \mathbf{x}^\top \nabla_{\mathbf{xx}}^2 \eta^{(n_1, n_2)}(\mu, \mathbf{x}) \Delta \mathbf{x}}. \quad (16)$$

Note that if  $(\mathbf{x}, \mathbf{s}; \mathbf{y}^{(1)}, \mathbf{s}_2^{(1)}; \dots; \mathbf{y}^{(K)}, \mathbf{s}_2^{(K)})$  is feasible for (14), then

$$(\mathbf{x}, \mathbf{s}; \mathbf{y}^{(1)}, \mathbf{s}_2^{(1)}; \dots; \mathbf{y}^{(K)}, \mathbf{s}_2^{(K)}) = (\mathbf{x}(\mu), \mathbf{s}(\mu); \mathbf{y}^{(1)}(\mu), \mathbf{s}_2^{(1)}(\mu); \dots; \mathbf{y}^{(K)}(\mu), \mathbf{s}_2^{(K)}(\mu)) \iff \delta(\mu, \mathbf{x}) = 0.$$

### 3 Relationships and computations

In this section, we first study the aspect of the relationship between the SLSIP problem and its discretization. Then, we obtain expressions for the derivatives of the recourse functions required in the rest of the paper.

### 3.1 Relationship of the SLSIP problem to the discretized problem

In this part, we show that the problem of finding an approximate minimizer of the SLSIP problem (5) and (6) can be reduced to that of finding an approximate solution:  $x \in \mathcal{F}$ ,  $\mathbf{y}^{(1)} \in \mathcal{F}_2^{(1)}(x), \dots, \mathbf{y}^{(K)} \in \mathcal{F}_2^{(K)}(x)$  of the discretized problem (7) and (8) under a certain condition on the measure of proximity  $\delta$  of the current point  $x$  to the central path.

We need the following lemma in Lemma 3.2 in order to be able to relate the approximate solutions of the SLSIP problem to the discretized problem.

**Lemma 3.1.** *Let  $(x, s; \mathbf{y}^{(1)}, s_2^{(1)}; \dots; \mathbf{y}^{(K)}, s^{(K)})$  be a feasible solution of the SLSIP problem (9) and (10) such that  $\delta \leq 1$ . Then the first- and second-stage dual multipliers  $\mathbf{v}$  and  $(\mathbf{v}^{(1)}; \dots; \mathbf{v}^{(K)})$  satisfy*

$$A^{(n_1)} \mathbf{v} + \sum_{k=1}^K T^{(k, n_2)} \mathbf{v}^{(k)} = \mathbf{c}, W^{(k, n_2)} \mathbf{v}^{(k)} = \mathbf{d}^{(k)}, \mathbf{v} \geq \mathbf{0}, \mathbf{v}^{(k)} \geq \mathbf{0}, k = 1, 2, \dots, K.$$

Moreover,

$$\mathbf{c}^\top \mathbf{x} - \mathbf{b}^{(n_1)\top} \mathbf{v} + \sum_{k=1}^K \left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} - \left( \mathbf{h}^{(k, n_2)} - T^{(k, n_2)\top} \mathbf{x} \right)^\top \mathbf{v}^{(k)} \right) \leq \mu(K+1) \left( n_1 + n_2 + \delta \sqrt{n_1 + n_2} \right).$$

*Proof.* The first statement follows by applying the first statement in [17, Theorem 2.4] to the DSILP problem (11) which is indeed equivalent to the SLSIP problem (9) and (10).

To prove the second statement, we now apply the second statement in [17, Theorem 2.4] to the subproblem (9) to get

$$\mathbf{c}^\top \mathbf{x} - \mathbf{b}^{(n_1)\top} \mathbf{v} \leq \mu \left( n_1 + \delta \sqrt{n_1} \right) \leq \mu \left( n_1 + n_2 + \delta \sqrt{n_1 + n_2} \right), \quad (17)$$

and to the subproblem (10) to get

$$\mathbf{d}^{(k)\top} \mathbf{y}^{(k)} - \left( \mathbf{h}^{(k, n_2)} - T^{(k, n_2)\top} \mathbf{x} \right)^\top \mathbf{v}^{(k)} \leq \mu \left( n_2 + \delta \sqrt{n_2} \right) \leq \mu \left( n_1 + n_2 + \delta \sqrt{n_1 + n_2} \right), k = 1, 2, \dots, K. \quad (18)$$

The result follows by summing over the right-hand side inequality in (17) and all the  $K$  right-hand side inequalities in (18).  $\square$

Lemma 3.1 demonstrates that if one can find  $(x, s, \mathbf{v}; \mathbf{y}^{(1)}, s_2^{(1)}, \mathbf{v}^{(1)}; \dots; \mathbf{y}^{(K)}, s^{(K)}, \mathbf{v}^{(K)})$  for which the point  $(x; \mathbf{y}^{(1)}; \dots; \mathbf{y}^{(K)})$  is feasible for the discretized problem (7) and (8) and if there holds  $\delta \leq 1$  with  $\mu \leq \epsilon / \left( (K+1) \left( n_1 + n_2 + \sqrt{n_1 + n_2} \right) \right)$ , then  $(x; \mathbf{y}^{(1)}; \dots; \mathbf{y}^{(K)})$  is an  $\epsilon$ -minimizer of the discretized problem (7) and (8). The following lemma takes one more step forward by demonstrating that  $(x; \mathbf{y}^{(1)}; \dots; \mathbf{y}^{(K)})$  is an  $\epsilon$ -minimizer of the original SLSIP problem (5) and (6) provided that  $(x; \mathbf{y}^{(1)}; \dots; \mathbf{y}^{(K)})$  satisfies the additional condition of being feasible for the problem (5) and (6). The idea of the proof of the following lemma is motivated by that of [8, Lemma 2.2].

**Lemma 3.2.** *Let  $\epsilon \in (0, 1)$  and let  $N_1 := \{1, 2, \dots, n_1\}$  and  $N_2 := \{1, 2, \dots, n_2\}$  be finite subsets. Suppose that  $\Gamma$  and  $\Lambda$  are compact, the mappings  $\gamma \rightarrow \mathbf{a}_\gamma$  and  $\gamma \rightarrow b_\gamma$  are continuous in  $\gamma$  on the set  $\Gamma$ , and the mappings  $\lambda \rightarrow \mathbf{w}_\lambda^{(k)}$ ,  $\lambda \rightarrow \mathbf{t}_\lambda^{(k)}$  and  $\lambda \rightarrow h_\lambda^{(k)}$  are continuous in  $\lambda$  on the set  $\Lambda$  for  $k = 1, 2, \dots, K$ . If  $x \in \mathcal{F}$ ,  $\mathbf{y}^{(1)} \in \mathcal{F}_2^{(1)}(x), \dots, \mathbf{y}^{(K)} \in \mathcal{F}_2^{(K)}(x)$  satisfy  $\delta \leq 1$  with  $\mu \leq \epsilon / \left( (K+1) \left( n_1 + n_2 + \sqrt{n_1 + n_2} \right) \right)$  is an  $\epsilon$ -minimizer of the SLSIP problem (5) and (6).*



*Proof.* By Lemma 3.1,  $(\mathbf{x}; \mathbf{y}^{(1)}; \dots; \mathbf{y}^{(K)})$  is an  $\epsilon$ -minimizer of the discretized problem (7) and (8) defined by the index sets  $N_1 := \{1, 2, \dots, n_1\}$  and  $N_2 := \{1, 2, \dots, n_2\}$ , and the dual multiplier  $(\mathbf{v}^{(n_1)}(\mu, \mathbf{x}); \mathbf{v}^{(1, n_2)}(\mu, \mathbf{x}); \dots; \mathbf{v}^{(K, n_2)}(\mu, \mathbf{x})) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \dots \times \mathbb{R}^{n_2}$  satisfies

$$A^{(n_1)} \mathbf{v}(\mu, \mathbf{x}) + \sum_{k=1}^K T^{(k, n_2)} \mathbf{v}^{(k)}(\mu, \mathbf{x}) = \mathbf{c}, \quad W^{(k, n_2)} \mathbf{v}^{(k)}(\mu, \mathbf{x}) = \mathbf{d}^{(k)}, \quad \mathbf{v}(\mu, \mathbf{x}) \geq \mathbf{0}, \quad \mathbf{v}^{(k)}(\mu, \mathbf{x}) \geq \mathbf{0}, \quad k = 1, 2, \dots, K.$$

Now, we consider any finite index sets  $Q_1 \subset \Gamma$  and  $Q_2 \subset \Lambda$  such that  $N_1 \subset Q_1$  and  $N_2 \subset Q_2$ . Let  $q_1$  and  $q_2$  be the cardinalities of  $Q_1$  and  $Q_2$ , respectively. We claim that  $(\mathbf{x}; \mathbf{y}^{(1)}; \dots; \mathbf{y}^{(K)})$  is an  $\epsilon$ -minimizer of the discretized problem (5) and (6) defined by the index sets  $Q_1$  and  $Q_2$ . To see this, we construct the dual multiplier  $(\bar{\mathbf{v}}^{(q_1)}(\mu, \mathbf{x}); \bar{\mathbf{v}}^{(1, q_2)}(\mu, \mathbf{x}); \dots; \bar{\mathbf{v}}^{(K, q_2)}(\mu, \mathbf{x})) \in \mathbb{R}^{q_1} \times \mathbb{R}^{q_2} \times \dots \times \mathbb{R}^{q_2}$  with

$$\bar{\mathbf{v}}^{(q_1)} := \begin{pmatrix} \mathbf{v}^{(n_1)}(\mu, \mathbf{x}) \\ \mathbf{0} \end{pmatrix} \quad \text{and} \quad \bar{\mathbf{v}}^{(k, q_2)} := \begin{pmatrix} \mathbf{v}^{(k, n_2)}(\mu, \mathbf{x}) \\ \mathbf{0} \end{pmatrix} \quad \text{for } k = 1, 2, \dots, K.$$

Recall that  $A^{(q_1)}$  (respectively,  $\mathbf{b}^{(q_1)}$ ) denotes a matrix (a vector) whose  $i^{\text{th}}$  column (component) is given by  $\mathbf{a}_i$  (respectively,  $b_i$ ),  $i \in Q_1$ . Similarly,  $T^{(k, q_2)}$  and  $W^{(k, q_2)}$  (respectively,  $\mathbf{h}^{(k, q_2)}$ ) denote(s) matrix(es) whose  $j^{\text{th}}$  column (component) is given by  $\mathbf{t}_j^{(k)}$  and  $\mathbf{w}_j^{(k)}$  (respectively,  $h_j^{(k)}$ ),  $j \in Q_2$ .

Because  $N_l \subset Q_l$ , for each  $l = 1, 2$ , the matrices  $A^{(q_1)}$ ,  $T^{(k, q_2)}$  and  $W^{(k, q_2)}$  contain  $A^{(n_1)}$ ,  $T^{(k, n_2)}$  and  $W^{(k, n_2)}$ , respectively, as submatrices. It follows that

$$\begin{aligned} A^{(q_1)} \bar{\mathbf{v}}^{(q_1)} + \sum_{k=1}^K T^{(k, q_2)} \bar{\mathbf{v}}^{(k, q_2)} &= A^{(n_1)} \mathbf{v}^{(n_1)}(\mu, \mathbf{x}) + \sum_{k=1}^K T^{(k, n_2)} \mathbf{v}^{(k, n_2)}(\mu, \mathbf{x}) = \mathbf{c}, \\ W^{(k, q_2)} \bar{\mathbf{v}}^{(k, q_2)} &= W^{(k, n_2)} \mathbf{v}^{(k, n_2)}(\mu, \mathbf{x}) = \mathbf{d}^{(k)}, \\ \bar{\mathbf{v}}^{(q_1)} \geq \mathbf{0}, \quad \bar{\mathbf{v}}^{(k, q_2)} \geq \mathbf{0}, \quad k &= 1, 2, \dots, K. \end{aligned}$$

This proves that  $(\bar{\mathbf{v}}^{(n_1)}; \bar{\mathbf{v}}^{(1, n_2)}; \dots; \bar{\mathbf{v}}^{(K, n_2)})$  is dual feasible. In addition, we have

$$\begin{aligned} \mathbf{c}^\top \mathbf{x} - \mathbf{b}^{(q_1)\top} \bar{\mathbf{v}}^{(q_1)} + \sum_{k=1}^K \left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} - \left( \mathbf{h}^{(k, q_2)} - T^{(k, q_2)\top} \mathbf{x} \right)^\top \bar{\mathbf{v}}^{(k, q_2)} \right) \\ = \mathbf{c}^\top \mathbf{x} - \mathbf{b}^{(n_1)\top} \mathbf{v} + \sum_{k=1}^K \left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} - \left( \mathbf{h}^{(k, n_2)} - T^{(k, n_2)\top} \mathbf{x} \right)^\top \mathbf{v}^{(k)} \right). \end{aligned}$$

Thus, the duality gap remains unchanged. By Lemma 3.1 again, we obtain

$$\begin{aligned} \varrho^{(q_1, q_2)} - \left( \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^K \left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} \right) \right) &= \varrho_1^{(q_1)} - \mathbf{c}^\top \mathbf{x} + \sum_{k=1}^K \varrho_2^{(k, q_2)} - \sum_{k=1}^K \left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} \right) \\ &\leq \mathbf{c}^\top \mathbf{x} - \mathbf{b}^{(q_1)\top} \bar{\mathbf{v}}^{(q_1)} + \sum_{k=1}^K \left( \mathbf{d}^{(k)\top} \mathbf{y}^{(k)} - \left( \mathbf{h}^{(k, q_2)} - T^{(k, q_2)\top} \mathbf{x} \right)^\top \bar{\mathbf{v}}^{(k, q_2)} \right) \\ &\leq (K+1) \left( n_1 + n_2 + \sqrt{n_1 + n_2} \right) \mu \\ &\leq \epsilon, \end{aligned}$$

where  $\varrho_1^{(q_1)}$  is the minimum value of the discretized subproblem  $\min_{\mathbf{x}} \{ \mathbf{c}^\top \mathbf{x} : \mathbf{a}_i^\top \mathbf{x} \geq b_i, 1 \leq i \leq n_1 \}$  defined by the index set  $Q_1$ ,  $\varrho_2^{(k, q_2)}$  is the minimum value of the discretized problem (8) defined by

the index set  $Q_2$ , for  $k = 1, 2, \dots, K$ , and  $\varrho^{(q_1, q_2)} := \varrho_1^{(q_1)} + \sum_{k=1}^K \varrho_2^{(k, q_2)}$  is the minimum value of the discretized problem (7) defined by the index sets  $Q_1$  and  $Q_2$ .

Recall that  $\varrho^{(k)}$  denotes the global minimum of the original problem (6). Let  $\varrho$  denote the global minimum of the original problem (5). Now, we are ready to apply a theorem of Gustafson [18] which concludes that under the compactness assumptions on  $\Gamma$  and  $\Lambda$  and the assumptions that the maps  $\gamma \rightarrow \mathbf{a}_\gamma, \gamma \rightarrow \mathbf{b}_\gamma, \lambda \rightarrow \mathbf{w}_\lambda^{(k)}, \lambda \rightarrow \mathbf{t}_\lambda^{(k)}$  and  $\lambda \rightarrow h_\lambda^{(k)}$  are continuous, the minimum value  $\varrho_2^{(k, q_2)}$  converges to  $\varrho^{(k)}$  for a certain sequence of index set  $Q_2$  with increasing cardinality, and the minimum value  $\varrho^{(q_1, q_2)}$  converges to  $\varrho$  for a certain sequence of index sets  $Q_1$  and  $Q_2$  with increasing cardinality.

Taking the limit along the sequences of index sets  $Q_1$  and  $Q_2$  in the above inequality, we get showing that  $(\mathbf{x}, \mathbf{y}^{(1)}, \dots, \mathbf{y}^{(K)})$  is indeed an  $\epsilon$ -minimizer of the SLSIP problem (5) and (6). The proof is complete.  $\square$

By Lemma 3.2, we can reduce the problem of finding an approximate minimizer of problem (5) and (6) to that of finding an approximate solution  $(\mathbf{x}, \mathbf{y}^{(1)}, \dots, \mathbf{y}^{(K)})$  of the discretized problem (9) and (10) such that  $\mathbf{x} \in \mathcal{F}, \mathbf{y}^{(1)} \in \mathcal{F}_2^{(1)}(\mathbf{x}), \dots, \mathbf{y}^{(K)} \in \mathcal{F}_2^{(K)}(\mathbf{x})$  and  $\delta(\mu, \mathbf{x}) \leq 1$ . So, in the proposed algorithms, we always maintain these two conditions, and we should be able to guarantee that we can successfully reduce the barrier parameter  $\mu$  and the algorithms terminate finitely. Resolving this is, in fact, the substance of Section 5.

### 3.2 Computation of $\nabla_x \eta(\mu, \mathbf{x})$ and $\nabla_{xx}^2 \eta(\mu, \mathbf{x})$

In this part, we compute the gradient and the Hessian of  $\eta^{(n_1, n_2)}(\mu, \mathbf{x})$ , which in turn requires obtaining a representation for the gradient and the Hessian of the barrier functions.

In order to compute the derivatives of  $\eta$  we need to determine the derivatives of the function  $\rho$ ,  $k = 1, 2, \dots, K$ . This requires computing the derivatives of  $\ell_1^{(n_1)}$  and  $v_1^{(n_1)}$  and the partial derivatives of  $\ell_2^{(k, n_2)}$  and  $v_2^{(k, n_2)}$  with respect to  $\mathbf{x}$  for  $k = 1, 2, \dots, K$ . Throughout the rest of this section and the next, we will drop the superscripts  $(k), (n_1)$  and  $(k, n_2)$  when it does not lead to confusion.

First, we compute the gradient and the Hessian of the logarithmic barriers  $\ell_1(\mathbf{x})$  and  $\ell_2(\mathbf{x}, \mathbf{y})$  with respect to  $\mathbf{x}$  and  $\mathbf{y}$ .

Note that

$$\begin{aligned} \nabla_x \mathbf{s}_1(\mathbf{x}) &= \nabla_x (A^\top \mathbf{x} - \mathbf{b}) = A^\top, \\ \nabla_x \mathbf{s}_2(\mathbf{x}, \mathbf{y}) &= \nabla_x (W^\top \mathbf{y} + T^\top \mathbf{x} - \mathbf{h}) = T^\top, \end{aligned}$$

and that

$$\begin{aligned} \nabla_x \mathbf{s}_1^{-1}(\mathbf{x}) &= -S_1^{-2} \nabla_x \mathbf{s}_1 = -S_1^{-2} A^\top, \\ \nabla_x \mathbf{s}_2^{-1}(\mathbf{x}, \mathbf{y}) &= -S_2^{-2} \nabla_x \mathbf{s}_2 = -S_2^{-2} T^\top. \end{aligned}$$

This implies that

$$\begin{aligned} \nabla_x \ell_1(\mathbf{x}) &= -(\nabla_x \mathbf{s}_1)^\top \mathbf{s}_1^{-1} = -A \mathbf{s}_1^{-1}, \\ \nabla_x \ell_2(\mathbf{x}, \mathbf{y}) &= -(\nabla_x \mathbf{s}_2)^\top \mathbf{s}_2^{-1} = -T \mathbf{s}_2^{-1}, \end{aligned}$$

and that

$$\begin{aligned} \nabla_{xx}^2 \ell_1(\mathbf{x}) &= -A \nabla_x \mathbf{s}_1^{-1} = A S_1^{-2} A^\top, \\ \nabla_{xx}^2 \ell_2(\mathbf{x}, \mathbf{y}) &= -T \nabla_x \mathbf{s}_2^{-1} = T S_2^{-2} T^\top. \end{aligned}$$

Note that the Hessian matrices are positive definite under Assumption 2.1 and the assumptions that  $\mathbf{s}_1 = \mathbf{s}_1(\mathbf{x}) > \mathbf{0}$  and  $\mathbf{s}_2 = \mathbf{s}_2(\mathbf{x}, \mathbf{y}) > \mathbf{0}$ .

Next, we compute the gradient and the Hessian of the volumetric barriers  $v_1(\mathbf{x})$  and  $v_2(\mathbf{x}, \mathbf{y})$  with respect to  $\mathbf{x}$ .

Throughout this section, we define

$$\begin{aligned} P_1 &:= P_1(\mathbf{s}_1) := S_1^{-1}A^\top(AS_1^{-2}A^\top)^{-1}AS_1^{-1}, \\ P_2 &:= P_2(\mathbf{s}_2) := S_2^{-1}W^\top(WS_2^{-2}W^\top)^{-1}WS_2^{-1}. \end{aligned}$$

Note that  $P_1$  and  $P_2$  act as the orthogonal projections onto the ranges of  $AS_1^{-1}$ ,  $TS_2^{-1}$  and  $WS_2^{-1}$ , respectively.

Let  $\sigma_1 := \sigma_1(\mathbf{s}_1)$  and  $\sigma_2 := \sigma_2(\mathbf{s}_2)$  denote the vectors equal to the diagonal of the projection matrices  $P_1$  and  $P_2$ , respectively. In other words,  $\sigma_{1i} = P_{1ii}$  and  $\sigma_{2j} = P_{2jj}$ , respectively, for  $i = 1, 2, \dots, n_1$  and  $j = 1, 2, \dots, n_2$ . Following our notations in Section 1, let  $\Sigma_1 := \text{diag}(\sigma_1)$  and  $\Sigma_2 := \text{diag}(\sigma_2)$ .

Derivations in the Appendix of Anstreicher [19] can be reconstructed and similar details contained therein can be adopted for our setting to obtain

$$\begin{aligned} \nabla_x v_1(\mathbf{x}) &= -AS_1^{-1}\sigma_1, \\ \nabla_x v_2(\mathbf{x}, \mathbf{y}) &= -TS_2^{-1}\sigma_2. \end{aligned} \tag{19}$$

and

$$\begin{aligned} \nabla_{xx}^2 v_1(\mathbf{x}) &= AS_1^{-1}(3\Sigma_1 - 2P_1 \odot P_1)S_1^{-1}A^\top, \\ \nabla_{xx}^2 v_2(\mathbf{x}, \mathbf{y}) &= TS_2^{-1}(3\Sigma_2 - 2P_2 \odot P_2)S_2^{-1}T^\top. \end{aligned} \tag{20}$$

We now compute the first and second order derivatives of  $\rho$  with respect to  $\mathbf{x}$ .

Define  $\varphi : \mathbb{R}^+ \times \mathcal{F}^\circ \times \mathcal{F}^{(k)}(\mathbf{x}) \rightarrow \mathbb{R}$  by

$$\varphi(\mu, \mathbf{x}, \mathbf{y}) := \mathbf{d}^\top \mathbf{y} + \mu \varsigma_2 v_2(\mathbf{x}, \mathbf{y}).$$

By (13) we then have

$$\rho(\mu, \mathbf{x}) = \min_{\mathbf{y} \in \mathcal{F}^{(k)}(\mathbf{x})} \varphi(\mu, \mathbf{x}, \mathbf{y})$$

and

$$\rho(\mu, \mathbf{x}) = \varphi(\mu, \mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} = \varphi(\mu, \mathbf{x}, \bar{\mathbf{y}}),$$

where  $\bar{\mathbf{y}}$  is the minimizer of (13). Observe that  $\bar{\mathbf{y}}$  is a function of  $\mathbf{x}$  and is defined by

$$\nabla_{\mathbf{y}} \varphi(\mu, \mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}. \tag{21}$$

Note that, by (21), we have  $\nabla_{\bar{\mathbf{y}}} v_2(\mathbf{x}, \bar{\mathbf{y}}) = \nabla_{\mathbf{y}} v_2(\mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} = -\frac{1}{\mu \varsigma_2} \mathbf{d}$ . This implies that

$$\nabla_{\bar{\mathbf{y}}\mathbf{x}}^2 v_2(\mathbf{x}, \bar{\mathbf{y}}) = \nabla_{\mathbf{x}} \nabla_{\bar{\mathbf{y}}} v_2(\mathbf{x}, \bar{\mathbf{y}}) = 0 \quad \text{and} \quad \nabla_{\bar{\mathbf{y}}\mathbf{x}\mathbf{x}}^3 v_2(\mathbf{x}, \bar{\mathbf{y}}) = \nabla_{\mathbf{x}\mathbf{x}}^2 \nabla_{\bar{\mathbf{y}}} v_2(\mathbf{x}, \bar{\mathbf{y}}) = 0.$$

We are now in a position to calculate the first and second order derivatives of  $\rho$  with respect to  $\mathbf{x}$ . We have

$$\begin{aligned} \nabla_x \rho(\mu, \mathbf{x}) &= [\nabla_x \varphi(\mu, \mathbf{x}, \mathbf{y}) + \nabla_{\mathbf{y}} \varphi(\mu, \mathbf{x}, \mathbf{y}) \nabla_x \mathbf{y}]|_{\mathbf{y}=\bar{\mathbf{y}}} \\ &= \nabla_x \varphi(\mu, \mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} + \nabla_{\mathbf{y}} \varphi(\mu, \mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} \nabla_x \mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} \\ &= \nabla_x \varphi(\mu, \mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} \\ &= \mu \varsigma_2 \nabla_x v_2(\mathbf{x}, \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} \\ &= \mu \varsigma_2 \nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}}), \end{aligned}$$

and

$$\begin{aligned} \nabla_{xx}^2 \rho(\mu, \mathbf{x}) &= \mu \nabla_x \{ \nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}}) \} \\ &= \mu \varsigma_2 \{ \nabla_{xx}^2 v_2(\mathbf{x}, \bar{\mathbf{y}}) + \nabla_{\bar{\mathbf{y}}\mathbf{x}}^2 v_{n_2}(\mathbf{x}, \bar{\mathbf{y}}) (\nabla_x \mathbf{y})|_{\mathbf{y}=\bar{\mathbf{y}}} \} \\ &= \mu \varsigma_2 \nabla_{xx}^2 v_2(\mathbf{x}, \bar{\mathbf{y}}). \end{aligned}$$

In summary, we obtain

$$\begin{aligned}\nabla_x \rho^{(k)}(\mu, \mathbf{x}) &= \mu \varsigma_2 \nabla_x v_2^{(k)}(\mathbf{x}, \bar{\mathbf{y}}^{(k)}), \\ \nabla_{xx}^2 \rho^{(k)}(\mu, \mathbf{x}) &= \mu \varsigma_2 \nabla_{xx}^2 v_2^{(k)}(\mathbf{x}, \bar{\mathbf{y}}^{(k)}),\end{aligned}\tag{22}$$

and therefore, by (14), we get

$$\begin{aligned}\nabla_x \eta(\mu, \mathbf{x}) &= \mathbf{c} + \mu \varsigma_1 \nabla_x v_1(\mathbf{x}) + \sum_{k=1}^K \mu \varsigma_2 \nabla_x v_2^{(k)}(\mathbf{x}, \bar{\mathbf{y}}^{(k)}), \\ \nabla_{xx}^2 \eta(\mu, \mathbf{x}) &= \mu \varsigma_1 \nabla_{xx}^2 v_1(\mathbf{x}) + \sum_{k=1}^K \mu \varsigma_2 \nabla_{xx}^2 v_2^{(k)}(\mathbf{x}, \bar{\mathbf{y}}^{(k)}),\end{aligned}\tag{23}$$

where  $\nabla_x v_1(\mathbf{x})$  and  $\nabla_x^2 v_2^{(k)}(\mathbf{x}, \bar{\mathbf{y}}^{(k)})$  are calculated in (19) and  $\nabla_{xx} v_1(\mathbf{x})$  and  $\nabla_{xx}^2 v_2^{(k)}(\mathbf{x}, \bar{\mathbf{y}}^{(k)})$  are calculated in (20).

## 4 Self-concordance properties of the volumetric barrier recourse

In this section, we prove that the recourse function with volumetric barrier is a strongly self-concordant function leading to a strongly self-concordant family with appropriate parameters. This allows us to develop volumetric barrier decomposition interior point algorithms for solving SLSIPs and establish their convergence and complexity analysis. As we mentioned in the introduction, our proofs in this section are different from those in Section 3 of [14], for which the authors heavily made use of their own results in [15, Section 4] after re-writing the underlying problem as a stochastic semidefinite programming problem. In comparison, although our proofs are not totally self-contained, the self-concordance results for the current setting are completely proven in the context of linear programming, which has the advantage of allowing very explicit and direct proofs.

### 4.1 Self-concordance of $\eta(\mu, \cdot)$

In this part, we prove that  $\eta(\mu, \cdot)$  is a strongly self-concordant barrier on  $\mathcal{F}^\circ$ . We have the following definition.

**Definition 4.1** (Nesterov and Nemirovskii [20, Definition 2.1.1]). Let  $E$  be a finite-dimensional real vector space,  $G$  be an open nonempty convex subset of  $E$ , and let  $g$  be a  $C^3$ , convex mapping from  $G$  to  $\mathbb{R}$ . Then  $g$  is called  $\alpha$ -self-concordant on  $G$  with the parameter  $\alpha > 0$  if for every  $\mathbf{x} \in G$  and  $\mathbf{h} \in E$ , the following inequality holds:

$$|\nabla_{xxx}^3 g(\mathbf{x})[\mathbf{h}, \mathbf{h}, \mathbf{h}]| \leq 2\alpha^{-1/2} (\nabla_{xx}^2 g(\mathbf{x})[\mathbf{h}, \mathbf{h}])^{3/2}.$$

An  $\alpha$ -self-concordant function  $g$  on  $G$  is called *strongly  $\alpha$ -self-concordant* if  $g$  tends to infinity for any sequence approaching a boundary point of  $G$ .

We note that in the above definition the set  $G$  is assumed to be open. However, relative openness would be sufficient to apply the definition. See also [20, Item A, Page 57].

Throughout this section, we define

$$\begin{aligned}Q &:= Q(\mathbf{x}, \mathbf{y}) := TS_2^{-1} \Sigma_2 S_2^{-1} T^\top, \\ \hat{\Delta} &:= \hat{\Delta} \mathbf{x} := -\left\{ \nabla_{xx}^2 v_2(\mathbf{x}, \mathbf{y}) \right\}^{-1} \nabla_x v_2(\mathbf{x}, \mathbf{y}), \\ \hat{\delta} &:= \hat{\delta}(\mathbf{x}, \mathbf{y}) := \sqrt{\hat{\Delta}^\top \nabla_{xx}^2 v_2(\mathbf{x}, \mathbf{y}) \hat{\Delta}}.\end{aligned}$$

The proof of self-concordance of  $\eta(\mu, \cdot)$  depends on the following three lemmas.

**Lemma 4.1.** For any  $(x, y)$  having  $s_2(x, y) > \mathbf{0}$ , we have  $0 \leq Q(x, y) \leq \nabla_{xx}^2 v_2(x, y) \leq 3Q(x, y)$ .

**Lemma 4.2.** For every  $p \in \mathbb{R}^{m_1}$ , we have  $\delta(x, y) \leq m_1^{1/4} \sqrt{Q(x, y)[p, p]} \leq m_1^{1/4} \sqrt{\nabla_{xx}^2 v_2(x, y)[p, p]}$ .

**Lemma 4.3.** Let  $\hat{x} := x + \hat{\Delta}x$  and  $\hat{y} := y(\hat{x})$  where  $\hat{\delta} < 0.1$ . Then for every  $p \in \mathbb{R}^{m_1}$ , we have

$$\left| \left( \nabla_{xx}^2 v_2(\hat{x}, \hat{y}) - \nabla_{xx}^2 v_2(x, y) \right) [p, p] \right| \leq \frac{30\hat{\delta}}{(1 - \hat{\delta})^4} Q(x, y)[p, p].$$

For proofs of Lemmas 4.1 and 4.2, similar details contained in the derivations of Equations (2.3) and (2.18) in [12] can be adopted and reconstructed for current setting. The proof of Lemma 4.3 is very similar to that of Theorem 2.3 in [12] except that our setting is based on stochasticity which brings the second-stage variable  $y$  besides  $x$ .

We can now state and prove the characterization the self concordance of  $v_1(\cdot)$  and  $v_2(\cdot, \cdot)$  for  $\mathcal{F}_1^\circ$  and  $\mathcal{F}_2^\circ$ , respectively. Then, we state and prove the following results which characterize the self-concordance of  $\rho(\mu, \cdot)$  and  $\eta(\mu, \cdot)$  for  $\mathcal{F}^\circ$ .

**Theorem 4.1.** Let  $\varsigma_1 = \varsigma_2 := 225\sqrt{m_1}$ . The functions  $\mu\varsigma_1 v_1(\cdot)$  and  $\mu\varsigma_2 v_2(\cdot, \cdot)$  are  $\mu$ -self-concordant barriers on  $\mathcal{F}_1^\circ$  and  $\mathcal{F}_2^\circ$ , respectively, for any fixed  $\mu > 0$ .

*Proof.* Let  $\mu > 0$  be fixed. Now, we verify that  $\mu\varsigma_2 v_2(\cdot, \cdot)$  is  $\mu$ -self-concordant on  $\mathcal{F}_2^\circ$ . Clearly, the map  $\mu\varsigma_2 v_2$  goes to infinity on any sequence approaching a boundary point of  $\mathcal{F}_2^\circ$ . By definition,

$$\nabla_{xxx}^3 v_2(x, y)[p, p, p] = \lim_{\tau \rightarrow 0} \frac{\nabla_{xx}^2 v_2(x + \tau p, y(x + \tau p))[p, p] - \nabla_{xx}^2 v_2(x, y)[p, p]}{\tau}$$

By using Lemmas 4.1, 4.2 and 4.3, with  $\hat{\Delta} = \tau p$ , we immediately get

$$\begin{aligned} \nabla_{xxx}^3 v_2(x, y)[p, p, p] &\leq \lim_{\tau \rightarrow 0} \frac{30\tau m_1^{1/4} \sqrt{\nabla_{xx}^2 v_2(x, y)[p, p]}}{\tau \left(1 - \tau m_1^{1/4} \sqrt{\nabla_{xx}^2 v_2(x, y)[p, p]}\right)^4} \nabla_{xx}^2 v_2(x, y)[p, p] \\ &= 30m_1^{1/4} \left(\nabla_{xx}^2 v_2(x, y)[p, p]\right)^{3/2}. \end{aligned} \quad (24)$$

The result follows by examining the effect on (24) when  $v_2(\cdot, \cdot)$  is multiplied by the factor  $\mu\varsigma_2$ . Thus  $\mu\varsigma_2 v_2(\cdot, \cdot)$  is  $\mu$ -self-concordant on  $\mathcal{F}_2^\circ$ . By applying similar argument on similar grounds, we can verify that  $\mu\varsigma_1 v_1(\cdot)$  is also  $\mu$ -self-concordant on  $\mathcal{F}_1^\circ$ . The proof is complete.  $\square$

**Theorem 4.2.** For any fixed  $\mu > 0$ ,  $\rho(\mu, \cdot)$  is  $\mu$ -self-concordant on  $\mathcal{F}^\circ$ , for  $k = 1, 2, \dots, K$ .

*Proof.* By using (22) we have

$$\begin{aligned} \nabla_{xxx}^3 \rho(\mu, x) &= \mu \varsigma_2 \nabla_x \{ \nabla_{xx}^2 v_2(x, \bar{y}) \} \\ &= \mu \varsigma_2 \{ \nabla_{xxx}^3 v_2(x, \bar{y}) + \nabla_{\bar{y}xx}^3 v_2(x, \bar{y}) (\nabla_x \bar{y})|_{y=\bar{y}} \} \\ &= \mu \varsigma_2 \nabla_{xxx}^3 v_2(x, \bar{y}). \end{aligned}$$

It follows that

$$\begin{aligned} \left| \nabla_{xxx}^3 \rho(\mu, x) [p, p, p] \right| &= \mu \varsigma_2 \left| \nabla_{xxx}^3 v_2(x, \bar{y}) [p, p, p] \right| \\ &\leq 30\mu \varsigma_2 m_1^{1/4} \left( \nabla_{xx}^2 v_2(x, \bar{y}) [p, p] \right)^{3/2} \\ &= 2\mu^{-1/2} \left( \mu \varsigma_2 \nabla_{xx}^2 v_2(x, \bar{y}) [p, p] \right)^{3/2} \\ &= 2\mu^{-1/2} \left( \nabla_{xx}^2 \rho(\mu, x) [p, p] \right)^{3/2}. \end{aligned}$$

This completes the proof.  $\square$

**Theorem 4.3.** For any fixed  $\mu > 0$ ,  $\eta(\mu, \cdot)$  is a  $\mu$ -self-concordant function on  $\mathcal{F}^\circ$ .

*Proof.* The theorem follows from Theorems 4.1 and 4.2 and [20, Proposition 2.1.1].  $\square$

## 4.2 Parameters of the self-concordant family

In this part, we verify that the family of functions  $\{\eta(\mu, \cdot) : \mu > 0\}$  is a strongly self-concordant family with appropriate parameters. First, we introduce the definition of self-concordant families.

**Definition 4.2** (Nesterov and Nemirovskii [20, Definition 3.1.1]). Let  $\mathbb{R}_{++}$  be the set of all positive real numbers. Let  $G$  be an open nonempty convex subset of  $\mathbb{R}^n$ . Let  $\mu \in \mathbb{R}_{++}$  and let  $g_\mu : \mathbb{R}_{++} \times G \rightarrow \mathbb{R}$  be a family of functions indexed by  $\mu$ . Let  $\alpha_1(\mu), \alpha_2(\mu), \alpha_3(\mu), \alpha_4(\mu)$  and  $\alpha_5(\mu) : \mathbb{R}_{++} \rightarrow \mathbb{R}_{++}$  be continuously differentiable functions on  $\mu$ . Then the family of functions  $\{g_\mu\}_{\mu \in \mathbb{R}_{++}}$  is called *strongly self-concordant with the parameters*  $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$ , if the following conditions hold:

- (i) The function  $g_\mu$  is continuous on  $\mathbb{R}_{++} \times G$ , and for fixed  $\mu \in \mathbb{R}_{++}$ ,  $g_\mu$  is convex on  $G$ . The function  $g_\mu$  has three partial derivatives on  $G$ , which are continuous on  $\mathbb{R}_{++} \times G$  and continuously differentiable with respect to  $\mu$  on  $\mathbb{R}_{++}$ .
- (ii) For any  $\mu \in \mathbb{R}_{++}$ , the function  $g_\mu$  is strongly  $\alpha_1(\mu)$ -self-concordant.
- (iii) For any  $(\mu, \mathbf{x}) \in \mathbb{R}_{++} \times G$  and any  $\mathbf{h} \in \mathbb{R}^n$ ,

$$\begin{aligned} |(\nabla_{\mathbf{x}} g_\mu(\mu, \mathbf{x})[\mathbf{h}])' - \{\ln \alpha_3(\mu)\}' \nabla_{\mathbf{x}} g_\mu(\mu, \mathbf{x})[\mathbf{h}]| &\leq \alpha_4(\mu) \alpha_1(\mu)^{\frac{1}{2}} \left( \nabla_{\mathbf{x}\mathbf{x}}^2 g_\mu(\mu, \mathbf{x})[\mathbf{h}, \mathbf{h}] \right)^{\frac{1}{2}}, \\ |(\nabla_{\mathbf{x}\mathbf{x}}^2 g_\mu(\mu, \mathbf{x})[\mathbf{h}, \mathbf{h}])' - \{\ln \alpha_2(\mu)\}' \nabla_{\mathbf{x}\mathbf{x}}^2 g_\mu(\mu, \mathbf{x})[\mathbf{h}, \mathbf{h}]| &\leq 2\alpha_5(\mu) \nabla_{\mathbf{x}\mathbf{x}}^2 g_\mu(\mu, \mathbf{x})[\mathbf{h}, \mathbf{h}]. \end{aligned}$$

The proof of self-concordancy of the family  $\{\eta(\mu, \cdot) : \mu > 0\}$  depends on the following two lemmas.

**Lemma 4.4.** For any  $\mu > 0$  and  $\mathbf{x} \in \mathcal{F}^\circ$ , we have

$$|(\nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x})'[\mathbf{p}, \mathbf{p}])| \leq \frac{1}{\mu} \nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x})[\mathbf{p}, \mathbf{p}], \quad \forall \mathbf{p} \in \mathbb{R}^{m_1}.$$

*Proof.* By differentiating  $\nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x})$  in (23) with respect to  $\mu$ , we get

$$\begin{aligned} \{\nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x})\}' &= \nabla_{\mathbf{x}\mathbf{x}}^2 v_1(\mathbf{x}) + \sum_{k=1}^K \left\{ \nabla_{\mathbf{x}\mathbf{x}}^2 v_2(\mathbf{x}, \bar{\mathbf{y}}) + \mu \nabla_{\mathbf{x}\mathbf{x}\bar{\mathbf{y}}}^3 v_2(\mathbf{x}, \bar{\mathbf{y}}) \cdot \bar{\mathbf{y}}' \right\} \\ &= \nabla_{\mathbf{x}\mathbf{x}}^2 v_1(\mathbf{x}) + \sum_{k=1}^K \nabla_{\mathbf{x}\mathbf{x}}^2 v_2(\mathbf{x}, \bar{\mathbf{y}}) = \frac{1}{\mu} \nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x}). \end{aligned}$$

Note that  $\nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x}) \geq 0$ , and therefore  $\frac{1}{\mu} \nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x})[\mathbf{p}, \mathbf{p}] \geq 0$  for any  $\mathbf{p} \in \mathbb{R}^{m_1}$ . The proof is complete.  $\square$

**Lemma 4.5.** For any  $\mu > 0$  and  $\mathbf{x} \in \mathcal{F}^\circ$ , we have

$$|\nabla_{\mathbf{x}} \eta'(\mu, \mathbf{x})^\top[\mathbf{p}]| \leq \sqrt{\frac{(m_1 \varsigma_1 + m_2 \varsigma_2)(1 + K)}{\mu}} \nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x})[\mathbf{p}, \mathbf{p}], \quad \forall \mathbf{p} \in \mathbb{R}^{m_1}.$$

*Proof.* Since  $P_2$  is a projection onto an  $m_2$ -dimensional space, we have  $P_2 = \mathbf{u}_1\mathbf{u}_1^\top + \mathbf{u}_2\mathbf{u}_2^\top + \dots + \mathbf{u}_{m_2}\mathbf{u}_{m_2}^\top$ , where  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{m_2}$  are orthonormal eigenvectors of  $P_2$  corresponding to the nonzero eigenvalues of  $P_2$ .

Using Lemma 4.1 we have  $\{\nabla_{xx}^2 v_2(\mathbf{x}, \mathbf{y})\}^{-1} \leq \{\mathbf{Q}(\mathbf{x}, \mathbf{y})\}^{-1}$ . This implies that

$$\begin{aligned} \nabla_x v_2(\mathbf{x}, \mathbf{y})^\top \{\nabla_{xx}^2 v_2(\mathbf{x}, \mathbf{y})\}^{-1} \nabla_x v_2(\mathbf{x}, \mathbf{y}) &\leq \nabla_x v_2(\mathbf{x}, \mathbf{y})^\top \{\mathbf{Q}(\mathbf{x}, \mathbf{y})\}^{-1} \nabla_x v_2(\mathbf{x}, \mathbf{y}) \\ &= \boldsymbol{\sigma}_2^\top S_2^{-1} T^\top \left( T S_2^{-1} \Sigma_2 S_2^{-1} T^\top \right)^{-1} T S_2^{-1} \boldsymbol{\sigma}_2 \\ &= \sqrt{\boldsymbol{\sigma}_2^\top} \left( \Sigma_2^{1/2} S_2^{-1} T^\top \left( T S_2^{-1} \Sigma_2 S_2^{-1} T^\top \right)^{-1} T S_2^{-1} \Sigma_2^{1/2} \right) \sqrt{\boldsymbol{\sigma}_2} \\ &\leq \sqrt{\boldsymbol{\sigma}_2^\top} \sqrt{\boldsymbol{\sigma}_2} \\ &= \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + \dots + \|\mathbf{u}_{m_2}\|^2 \\ &= m_2. \end{aligned}$$

The above bound is equivalent to

$$|\nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}})[\mathbf{p}]| \leq \sqrt{m_2 \nabla_{xx}^2 v_2(\mathbf{x}, \bar{\mathbf{y}})[\mathbf{p}, \mathbf{p}]}, \quad \forall \mathbf{p} \in \mathbb{R}^{m_1}. \quad (25)$$

Similarly, we can show that

$$|\nabla_x v_1(\mathbf{x})[\mathbf{p}]| \leq \sqrt{m_1 \nabla_{xx}^2 v_1(\mathbf{x})[\mathbf{p}, \mathbf{p}]}, \quad \forall \mathbf{p} \in \mathbb{R}^{m_1}. \quad (26)$$

By differentiating  $\nabla_x \eta(\mu, \mathbf{x})$  in (23) with respect to  $\mu$ , we have

$$\begin{aligned} \nabla_x \eta'(\mu, \mathbf{x}) &= \varsigma_1 \nabla_x v_1(\mathbf{x}) + \sum_{k=1}^K \varsigma_2 \{ \nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}}) + \mu \nabla_{x\bar{\mathbf{y}}}^2 v_2(\mathbf{x}, \bar{\mathbf{y}}) \cdot (\bar{\mathbf{y}})' \} \\ &= \varsigma_1 \nabla_x v_1(\mathbf{x}) + \sum_{k=1}^K \varsigma_2 \nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}}). \end{aligned}$$

It follows that, for all  $\mathbf{p} \in \mathbb{R}^{m_1}$ , by using (25) and (26) we obtain

$$\begin{aligned} |\nabla_x \eta'(\mu, \mathbf{x})[\mathbf{p}]| &= \left| \left( \varsigma_1 \nabla_x v_1(\mathbf{x}) + \sum_{k=1}^K \varsigma_2 \nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}}) \right) [\mathbf{p}] \right| \\ &\leq |\varsigma_1 \nabla_x v_1(\mathbf{x})[\mathbf{p}]| + \sum_{k=1}^K |\varsigma_2 \nabla_x v_2(\mathbf{x}, \bar{\mathbf{y}})[\mathbf{p}]| \\ &\leq \sqrt{m_1 \varsigma_1^2 \nabla_{xx}^2 v_1(\mathbf{x})[\mathbf{p}, \mathbf{p}]} + \sum_{k=1}^K \sqrt{m_2 \varsigma_2^2 \nabla_{xx}^2 v_2(\mathbf{x}, \bar{\mathbf{y}})[\mathbf{p}, \mathbf{p}]} \\ &\leq \sqrt{(m_1 \varsigma_1) \varsigma_1 \nabla_{xx}^2 v_1(\mathbf{x})[\mathbf{p}, \mathbf{p}]} + \sum_{k=1}^K \sqrt{(m_2 \varsigma_2) \varsigma_2 \nabla_{xx}^2 v_2(\mathbf{x}, \bar{\mathbf{y}})[\mathbf{p}, \mathbf{p}]} \\ &\leq \sqrt{(1+K)(m_1 \varsigma_1 + m_2 \varsigma_2) \left( \varsigma_1 \nabla_{xx}^2 v_1(\mathbf{x})[\mathbf{p}, \mathbf{p}] + \sum_{k=1}^K \varsigma_2 \nabla_{xx}^2 v_2(\mathbf{x}, \bar{\mathbf{y}})[\mathbf{p}, \mathbf{p}] \right)} \\ &= \sqrt{\frac{(1+K)(m_1 \varsigma_1 + m_2 \varsigma_2)}{\mu} \nabla_{xx}^2 \eta(\mu, \mathbf{x})[\mathbf{p}, \mathbf{p}]}. \end{aligned}$$

The proof is complete.  $\square$

**Theorem 4.4.** *The family  $\{\eta(\mu, \cdot) : \mu > 0\}$  is a strongly self-concordant family with the following parameters*

$$\alpha_1(\mu) = \mu, \alpha_2(\mu) = \alpha_3(\mu) = 1, \alpha_4(\mu) = \frac{\sqrt{(1+K)(m_1\varsigma_1 + m_2\varsigma_2)}}{\mu}, \alpha_5(\mu) = \frac{1}{2\mu}.$$

*Proof.* It can be easily seen that condition (i) of Definition 4.2 is satisfied. Theorem 4.3 shows that condition (ii) holds and Lemmas 4.4 and 4.5 show that condition (iii) holds.  $\square$

## 5 Volumetric barrier decomposition algorithms and complexity

Based on the self-concordance analysis established in Section 4, we develop a volumetric barrier decomposition algorithm for SLSIPs, which is formally stated in Algorithm 5.1.

---

**Algorithm 5.1.** Volumetric barrier decomposition algorithm for SLSIP

---

### Begin algorithm

- 1: **Initialize**  $\epsilon, \gamma, \theta, \beta, \mu^0, \mathbf{x}^0$
- 2:     Initial set of  $2m_1$  constraints  $A^{(2m_1)\top} \mathbf{x} \geq \mathbf{b}^{(2m_1)}$
- 3:     Initial set of  $2m_2$  constraints  $W^{(k, 2m_2)\top} \mathbf{y}^{(k)} \geq \mathbf{h}^{(k, 2m_2)} - T^{(k, 2m_2)\top} \mathbf{x}, k = 1, 2, \dots, K$
- Ensure:**  $\epsilon > 0, \gamma \in (0, 1), \theta > 0, \beta > 0, \mu^0 > 0, \mathbf{x}^0 \in \mathcal{F}^\circ$
- 4:     Set  $n_1 := 2m_1, n_2 := 2m_2, \mu := \mu^0, \mathbf{x} := \mathbf{x}^0$
- 5:     **While**  $\mu \geq \epsilon / ((K+1)(n_1 + n_2 + \sqrt{n_1 + n_2}))$  **do**
- 6:         **For**  $k = 1, 2, \dots, K$  **do**
- 7:             Solve (10) to obtain  $\bar{\mathbf{y}}^{(k)}$
- 8:         **End for**
- 9:         Compute  $\Delta \mathbf{x}$  using (23) and (15)
- 10:         Compute  $\delta(\mu, \mathbf{x})$  using (23) and (16)
- 11:         Set  $\mathbf{x} := \mathbf{x} + \theta \Delta \mathbf{x}$
- 12:          $\ddagger$  Add a constraint
- 13:         **If**  $\exists i \in \Gamma$  such that  $s_i < 0$  **then**
- 14:             Compute  $\xi_i := \sqrt{\varsigma_1 \mathbf{a}_i^\top \{\nabla_{xx}^2 v_1^{(n_1)}(\mathbf{x})\}^{-1} \mathbf{a}_i}, \tilde{b}_i := \mathbf{a}_i^\top \mathbf{x} - \beta \xi_i$
- 15:             Set  $A^{(n_1+1)} := (A^{(n_1)}, \mathbf{a}_i), \mathbf{b}^{(n_1+1)} := \begin{pmatrix} \mathbf{b}^{(n_1)} \\ \tilde{b}_i \end{pmatrix}$
- 16:             Set  $n_1 := n_1 + 1$
- 17:             **While**  $\delta > \beta$  **do**
- 18:                 **For**  $k = 1, 2, \dots, K$  **do**
- 20:                     Solve (10) to obtain  $\bar{\mathbf{y}}^{(k)}$
- 21:                 **End for**
- 22:                 Compute  $\Delta \mathbf{x}$  using (23) and (15)
- 23:                 Compute  $\delta(\mu, \mathbf{x})$  using (23) and (16)
- 24:                 Set  $\mathbf{x} := \mathbf{x} + \theta \Delta \mathbf{x}$
- 25:             **End while**
- 26:         **End if**
- 27:         **For**  $k = 1, \dots, K$  **do**
- 28:              $\ddagger$  Add a constraint
- 29:             **If**  $\exists j \in \Lambda$  such that  $s_j^{(k)} < 0$  **then**



```

30: Compute  $\zeta_j^{(k)} := \sqrt{\varsigma_2 \mathbf{w}_j^{(k)\top} \{\nabla_{xx}^2 v_2^{(k,n_2)}(\mathbf{x}, \mathbf{y}^{(k)})\}^{-1} \mathbf{w}_j^{(k)}, \tilde{h}_j^{(k)} := \mathbf{w}_j^{(k)\top} \mathbf{y}^{(k)} + \tilde{\mathbf{t}}_j^{(k)\top} \mathbf{x} - \beta \zeta_j^{(k)}$ 
31: Set  $W^{(k,n_2+1)} := (W^{(k,n_2)}, \mathbf{w}_j^{(k)})$ ,  $T^{(k,n_2+1)} := (T^{(k,n_2)}, \mathbf{t}_j^{(k)})$ ,  $\mathbf{h}^{(k,n_2+1)} := \begin{pmatrix} \mathbf{h}^{(k,n_2)} \\ \tilde{h}_j^{(k)} \end{pmatrix}$ 
32: Set  $(k, n_2) := (k, n_2 + 1)$ 
33: While  $\delta > \beta$  do
34:     For  $k = 1, 2, \dots, K$  do
35:         Solve (10) to obtain  $\bar{\mathbf{y}}^{(k)}$ 
36:     End for
37:     Compute  $\Delta \mathbf{x}$  using (23) and (15)
38:     Compute  $\delta(\mu, \mathbf{x})$  using (23) and (16)
39:     Set  $\mathbf{x} := \mathbf{x} + \theta \Delta \mathbf{x}$ 
40: End while
41: End if
42: End for
43:      $\mu := \gamma \mu$ 
44: End while
End algorithm

```

---

In Algorithm 5.1, we use  $\mu = \mu_0$  as the starting value for the barrier parameter,  $\epsilon$  as the desired accuracy of the final solution, and  $\gamma$  as the reduction parameter. We also use  $\beta$  as a threshold for the measure of the proximity  $\delta$  of the current point  $\mathbf{x}$  to the central path. We start with  $\mathbf{x}^0$  as a given first-stage interior point (possibly infeasible) that satisfies the initial set of  $2m_1$  constraints  $A^{(2m_1)\top} \mathbf{x} \geq \mathbf{b}^{(2m_1)}$  with an initial constraint matrix  $A^{(2m_1)}$  and an initial right-hand side vector  $\mathbf{b}^{(2m_1)}$ . We obtain second-stage interior points (possibly infeasible)  $\bar{\mathbf{y}}^{(1)}, \bar{\mathbf{y}}^{(2)}, \dots, \bar{\mathbf{y}}^{(K)}$  by solving Problem (10) projected to the initial set of  $2m_2$  constraints  $W^{(k,2m_2)\top} \mathbf{y}^{(k)} \geq \mathbf{h}^{(k,2m_2)} - T^{(k,2m_2)\top} \mathbf{x}$  with initial constraint matrices  $W^{(k,2m_2)}$  and  $T^{(k,2m_2)}$  and an initial right-hand side vector  $\mathbf{h}^{(k,2m_2)}$  for  $k = 1, 2, \dots, K$ .

For convenience, because it is our assumption that  $\mathbf{x}^0 \in \mathcal{F}_1^\circ \subseteq [0, 1]^{m_1}$ , we consider the set  $\{0 \leq x_i^0 \leq 1, i = 1, 2, \dots, m_1\}$  as our initial set of  $2m_1$  constraints; these constraints can be written as  $A^{(2m_1)\top} \mathbf{x} \geq \mathbf{b}^{(2m_1)}$ . Similarly, because it is our assumption that  $\bar{\mathbf{y}}^{(k)} \in \mathcal{F}_2^{\circ(k)} \subseteq [0, 1]^{m_2}$ , we consider the set  $\{0 \leq \bar{y}_i^{(k)} \leq 1, i = 1, 2, \dots, m_2\}$  as our initial set of  $2m_2$  constraints; these constraints can be written as  $W^{(k,2m_2)\top} \mathbf{y}^{(k)} \geq \mathbf{h}^{(k,2m_2)} - T^{(k,2m_2)\top} \mathbf{x}$ , for  $k = 1, 2, \dots, K$ .

Impacted by the manner of selecting the parameter  $\gamma$  in Algorithm 5.1, we have two variants of algorithms: The short-step algorithm and the long-step algorithm. Below we identify suitable values for the algorithmic parameters  $\gamma$  and  $\beta$  introduced in Algorithm 5.1.

$$\begin{aligned}
\gamma &:= \begin{cases} 1 - \frac{0.1}{\sqrt{(1+K)(m_1\varsigma_1 + m_2\varsigma_2)}}, & \text{if the short-step algorithm is performed;} \\ \text{an arbitrary value in } (0, 1), & \text{if the long-step algorithm is performed,} \end{cases} \\
\beta &:= \begin{cases} (2 - \sqrt{3})/2, & \text{if the short-step algorithm is performed;} \\ 1/6, & \text{if the long-step algorithm is performed.} \end{cases}
\end{aligned} \tag{27}$$

Note that if the current point  $\mathbf{x}$  is too far away from the central path in the sense that  $\delta > \beta$ , Newton's method is applied to find a point close to the central path, then the value of  $\mu$  is reduced by

a factor  $\gamma$  and the whole process is repeated until the value of  $\mu$  is within the tolerance  $\epsilon$ . Algorithm 5.1 approximately traces the central path as  $\mu$  approaches zero. This ends up in a strictly feasible  $\epsilon$ -optimal solution of the problem. Note also that when an iterate becomes infeasible, a new cut is introduced and the algorithm attempts to move to a new central point.

Theorems 5.1 and 5.2 present the complexity analysis for the short-step algorithm and the long-step algorithm, respectively.

In the short-step algorithm, the barrier parameter in each iteration is decreased by the factor  $\gamma$  given in (27). The  $k^{\text{th}}$  iteration of the short-step algorithm is performed as follows: At the beginning of the iteration, we have  $\mu^{(k-1)}$  and  $\mathbf{x}^{(k-1)}$  on hand and  $\mathbf{x}^{(k-1)}$  is close to the center path, i.e.,  $\delta(\mu^{(k-1)}, \mathbf{x}^{(k-1)}) \leq \beta$ . After we reduce the barrier parameter  $\mu$  from  $\mu^{(k-1)}$  to  $\mu^k := \gamma\mu^{(k-1)}$ , we have that  $\delta(\mu^k, \mathbf{x}^{(k-1)}) \leq 2\beta$ . Then we take a full Newton step with size  $\theta = 1$  to produce a new point  $\mathbf{x}^k$  with  $\delta(\mu^k, \mathbf{x}^k) \leq \beta$ . The following theorem presents the complexity result for short-step algorithm.

**Theorem 5.1.** *Assume that the maximum numbers of cuts generated by Algorithm 5.1 are finite and are denoted by  $n_{1\max}$  and  $n_{2\max}$  in the first- and second-stage problems, respectively. If the starting point  $\mathbf{x}^0$  is sufficiently close to the central path, i.e.,  $\delta(\mu^0, \mathbf{x}^0) \leq \beta$ , then the short-step algorithm reduces the barrier parameter  $\mu$  at a linear rate and terminates with at most*

$$\mathcal{O}\left(\sqrt{(1+K)(m_1 \varsigma_1 + m_2 \varsigma_2)} \ln\left(\frac{(1+K)(n_{1\max} + n_{2\max}) \mu^0}{\epsilon}\right)\right)$$

iterations.

*Proof.* See Appendix A.1. □

In the long-step algorithm, we decrease the barrier parameter by an arbitrary constant factor  $\gamma \in (0, 1)$ . It has a potential for larger decrease on the objective function value, however, several damped Newton steps might be needed for restoring the proximity to the central path. The  $k^{\text{th}}$  iteration of the long-step algorithm is performed as follows: At the beginning of the iteration we have a point  $\mathbf{x}^{(k-1)}$ , which is sufficiently close to  $\mathbf{x}(\mu^{(k-1)})$ , where  $\mathbf{x}(\mu^{(k-1)})$  is the solution to (12) for  $\mu := \mu^{(k-1)}$ . We reduce the barrier parameter from  $\mu^{(k-1)}$  to  $\mu^k := \gamma\mu^{(k-1)}$ , where  $\gamma \in (0, 1)$ , and then start the searching to find a point  $\mathbf{x}^k$  that is sufficiently close to  $\mathbf{x}(\mu^k)$ . The long-step algorithm generates a finite sequence consisting of  $N$  points in  $\mathcal{F}_0$ , and we finally take  $\mathbf{x}^k$  to be equal to the last point of this sequence. The following theorem presents the complexity result for long-step algorithm.

**Theorem 5.2.** *Assume that the maximum numbers of cuts generated by Algorithm 5.1 are finite and are denoted by  $n_{1\max}$  and  $n_{2\max}$  in the first- and second-stage problems, respectively. If the starting point  $\mathbf{x}^0$  is sufficiently close to the central path, i.e.,  $\delta(\mu^0, \mathbf{x}^0) \leq \beta$ , then the long-step algorithm reduces the barrier parameter  $\mu$  at a linear rate and terminates with at most*

$$\mathcal{O}\left((1+K)(m_1 \varsigma_1 + m_2 \varsigma_2) \ln\left(\frac{(1+K)(n_{1\max} + n_{2\max}) \mu^0}{\epsilon}\right)\right)$$

iterations.

*Proof.* See Appendix A.2. □

It is clear that the dominant terms in the complexity expressions in Theorems 5.1 and 5.2 are given in terms of the problem dimension and the number of realizations, and, most notably, they

are not given in terms of maximum numbers of cuts to be generated. In fact, in case the logarithmic barriers  $\ell_1$  and  $\ell_2$  are used instead of the volumetric barriers  $v_1$  and  $v_2$  in the first- and second-stage problems, respectively, it can be shown that the complexity expressions in Theorems 5.1 and 5.2 become

$$\mathcal{O}\left(\sqrt{n_{1\max} + Kn_{2\max}} \ln\left(\frac{(n_{1\max} + Kn_{2\max}) \mu^0}{\epsilon}\right)\right)$$

and

$$\mathcal{O}\left((n_{1\max} + Kn_{2\max}) \ln\left(\frac{(n_{1\max} + Kn_{2\max}) \mu^0}{\epsilon}\right)\right),$$

respectively, which have more complexity than those obtained in Theorems 5.1 and 5.2 because of contributing  $n_{1\max}$  and  $n_{1\max}$  in the leading terms and bounding them is generally difficult.

Note that the complexity results in Theorems 5.1 and 5.2 are the counterparts of those in Theorems 6 and 7 in [14] for two-stage stochastic quadratic linear programs with recourse, those in Theorems 3 and 4 in [16] for two-stage stochastic second-order programs with recourse, and those in Theorems 4 and 5 in [15] for two-stage stochastic semidefinite programs with recourse. Note also that the “rich flavor” hidden inside the volumetric barrier can be tasted in Theorems 5.1 and 5.2 more than in their counterpart theorems in [14–16]. The reason for this is that there are no cuts to be generated in the optimization problems studied in [14–16], which in turns makes no big difference by replacing  $m_1$  and  $m_2$  with  $n_1$  and  $n_2$  in case the volumetric barrier is not used in [14–16].

Since the maximum numbers of cuts to be generated contribute in the log-terms in the complexity expressions, we also bound the numbers  $n_{1\max}$  and  $n_{2\max}$ . Luo et al. [8] proved the existence of an upper bound on the number of cuts to be generated by the short-step algorithm. Taking into account such a bound in [8] and applying this to our problem setting, we conclude the following result which bounds the number of cuts to be generated by the short-step algorithm. The following theorem follows directly from the definition of self-concordance and [8, Theorem 4.2].

**Theorem 5.3.** *Assume the hypothesis of Lemma 3.2 holds. Let  $r_1$  be the radius of the largest Euclidean ball contained in  $\mathcal{F}_1$  and  $r_2$  be the radius of the largest Euclidean ball contained in  $\mathcal{F}_2^{(k)}$  for each  $k = 1, 2, \dots, K$ . Assume also that the short-step algorithm is used in Algorithm 5.1. Then for  $\epsilon > 0$ , Algorithm 5.1 terminates with an  $\epsilon$ -minimizer of Problem (5) and (6) after generating a total of at most*

$$\mathcal{O}^*\left(\frac{m_1^2}{r_1^2} e^4 \frac{\sqrt{m_1}}{\epsilon}\right)$$

*cuts for the first-stage subproblem, and generating a total of at most*

$$\mathcal{O}^*\left(\frac{m_2^2}{r_2^2} e^4 \frac{\sqrt{m_2}}{\epsilon}\right)$$

*cuts for the second-stage subproblem, where the superscript “\*” indicates that lower order terms have been ignored.*

We conclude from Theorem 5.3 that the log-terms in the complexity expressions Theorems 5.1 and 5.2 are given in terms of the problem dimension and the radii of the largest Euclidean balls contained in the feasibility sets. On the other side, we also conclude from the above discussion that the dominant complexity terms are not affected by the radii of the largest Euclidean balls contained in the feasibility sets.

## 6 Conclusion

In this paper, we have studied the two-stage stochastic linear semi-infinite programming problem. We have develop and analyzed volumetric barrier decomposition interior point algorithms for solving this class of optimization problems. We have proved the convergence results by showing that the volumetric barrier associated with the recourse function forms a self-concordant family. We have also described and analyzed short- and long-step variant of the algorithm that follow the primal central trajectory of the first-stage problem.

We have seen that, for a stochastic linear semi-infinite program with dimensions  $m_1$  and  $m_2$  in the first- and second-stage problems, respectively, and for  $K$  number of realizations, when the maximum numbers of cuts generated are  $n_{1\max}$  and  $n_{2\max}$  in the first- and second-stage problems, respectively, we need at most  $O\left(\sqrt{(1+K)(m_1 \varsigma_1 + m_2 \varsigma_2)} \ln((1+K)(n_{1\max} + n_{2\max}) \mu^0/\epsilon)\right)$  Newton iterations in the short-step algorithm to follow the central path from a starting value of the barrier parameter  $\mu^0$  to a terminating value  $\epsilon/((K+1)(n_{1\max} + n_{2\max} + \sqrt{n_{1\max} + n_{2\max}}))$ , and we need at most  $O\left((1+K)(m_1 \varsigma_1 + m_2 \varsigma_2) \ln((1+K)(n_{1\max} + n_{2\max}) \mu^0/\epsilon)\right)$  Newton iterations in the long-step algorithm for this recentering, where  $\epsilon$  is the desired accuracy of the final solution. Note that the dominant terms in the above complexity expressions are given in terms of the problem dimension and the number of realizations, and, most notably, they are not given in terms of maximum numbers of cuts to be generated, which means that they are not affected by the radii of the largest Euclidean balls contained in the feasibility sets.

Future work is devoted to develop a cutting surface algorithm for solving two-stage stochastic convex semi-infinite programming with recourse. The idea would be to work under a setting where cuts more general than linear can be used.

**Acknowledgment** A substantial part of this work was performed while the first author was visiting the Center for Applied and Computational Mathematics at Rochester Institute of Technology, NY, USA. The authors would like to thank Ayat Ababneh from The Ohio State University for reading the paper and providing valuable comments especially those on Subsection 3.2.

## 7 Appendix A. Complexity proofs

In this appendix, we present proofs for the complexity results stated in Section 5 that bound the number of iterations. The general scheme of our proofs follows the lines of the proofs from [16] and [15]. The proof of Theorem 5.1 for the short-step algorithm is given in Appendix A.1, and the proof of Theorem 5.2 for the long-step algorithm is given in Appendix A.2. Throughout this appendix, we will drop the superscripts  $(k)$ ,  $(n_1)$  and  $(k, n_2)$  when it does not lead to confusion.

The proofs make use of the following proposition which is due to [20, Theorem 2.1.1].

**Proposition 7.1.** *For any  $\mu > 0$  and  $x \in \mathcal{F}_0$ . Then for  $\delta < 1, \tau \in [0, 1]$  and any  $p \in \mathbb{R}^{m_1}$  we have*

$$\nabla_{xx}^2 \eta(\mu, x + \tau \Delta x)[p, p] \leq (1 - \tau \delta)^{-2} \nabla_{xx}^2 \eta(\mu, x)[p, p].$$

The following lemma describes the behavior of the Newton method as applied to  $\eta(\mu, \cdot)$ . This lemma is essentially [20, Theorem 2.2.3].

**Lemma 7.1.** *For any  $\mu > 0$  and  $x \in \mathcal{F}_0$ , let  $x^+ := x + \Delta x$ ,  $\Delta x^+$  be the Newton direction calculated at  $x^+$ , and  $\delta(\mu, x^+) := \sqrt{\frac{1}{\mu} \nabla_{xx}^2 \eta(\mu, x^+)[\Delta x^+, \Delta x^+]}$ . Then the following relations hold:*

(i) If  $\delta < 2 - \sqrt{3}$ , then  $\delta(\mu, \mathbf{x}^+) \leq \left(\frac{\delta}{1-\delta}\right)^2 \leq \frac{\delta}{2}$ .

(ii) If  $\delta \geq 2 - \sqrt{3}$ , then  $\eta(\mu, \mathbf{x}) - \eta(\mu, \mathbf{x} + \bar{\theta}\Delta\mathbf{x}) \geq \mu(\delta - \ln(1 + \delta))$ , where  $\bar{\theta} = (1 + \delta)^{-1}$ .

### A.1. Complexity proof of the short-step algorithm

The proof of Theorem 5.1 makes use of the following proposition which is a restatement of [20, Theorem 3.1.1].

**Proposition 7.2.** Let  $\chi_\kappa(\eta; \mu, \mu^+) := \left(1 + \frac{\sqrt{(1+K)(m_1\varsigma_1 + m_2\varsigma_2)}}{\kappa}\right) \ln\gamma^{-1}$ . Assume that  $\delta(\mu, \mathbf{x}) < \kappa$  and  $\mu^+ := \gamma\mu$  satisfies

$$\chi_\kappa(\eta; \mu, \mu^+) \leq 1 - \frac{\delta(\mu, \mathbf{x})}{\kappa}.$$

Then  $\delta(\mu^+, \mathbf{x}) < \kappa$ .

We also need the following lemma.

**Lemma 7.2.** Let  $\mu^+ = \gamma\mu$ . Then  $\delta(\mu^+, \mathbf{x}) \leq 2\beta$ .

*Proof.* Let  $\kappa := 2\beta = 2 - \sqrt{3}$ . Since  $\delta(\mu, \mathbf{x}) \leq \kappa/2$ , one can verify that for  $\sigma \leq 0.1$ ,  $\mu^+$  satisfies

$$\chi_\kappa(\eta; \mu, \mu^+) \leq \frac{1}{2} \leq 1 - \frac{\delta(\mu, \mathbf{x})}{\kappa}.$$

By Proposition 7.2, we have  $\delta(\mu^+, \mathbf{x}) \leq \kappa$ . □

From Lemmas 7.1(i) and 7.2, we deduce that we can reduce the parameter  $\mu$  by the factor  $\gamma$  given in (27), at each iteration, and that only one Newton step is sufficient to restore proximity to the central path. Hence, Theorem 5.1 follows.

### A.2. Complexity proof of the long-step algorithm

For  $\mathbf{x} \in \mathcal{F}_0$  and  $\mu > 0$ , we define the function  $\phi(\mu, \mathbf{x}) := \eta(\mu, \mathbf{x}(\mu)) - \eta(\mu, \mathbf{x})$ , which represents the difference between the objective value  $\eta(\mu^k, \mathbf{x}^{(k)})$  at the end of  $k^{\text{th}}$  iteration and the minimum objective value  $\eta(\mu^k, \mathbf{x}(\mu^{k-1}))$  at the beginning of  $k^{\text{th}}$  iteration. Then the task is to find an upper bound on  $\phi(\mu, \mathbf{x})$ . To do so, we first give upper bounds on  $\phi(\mu, \mathbf{x})$  and  $\phi'(\mu, \mathbf{x})$  respectively.

The proof of Theorem 5.2 makes use of the following lemma, in which its proof is quite similar to that of [15, Lemma 6].

**Lemma 7.3.** Let  $\mu > 0$  and  $\mathbf{x} \in \mathcal{F}_0$ , we denote  $\tilde{\Delta}\mathbf{x} := \mathbf{x} - \mathbf{x}(\mu)$  and define

$$\tilde{\delta} := \tilde{\delta}(\mu, \mathbf{x}) = \sqrt{\frac{1}{\mu} \nabla_{\mathbf{x}\mathbf{x}}^2 \eta(\mu, \mathbf{x}) [\tilde{\Delta}\mathbf{x}, \tilde{\Delta}\mathbf{x}]}.$$

For any  $\mu > 0$  and  $\mathbf{x} \in \mathcal{F}_0$ , if  $\tilde{\delta} < 1$ , then the following inequalities hold:

$$\begin{aligned} \phi(\mu, \mathbf{x}) &\leq \mu \left( \frac{\tilde{\delta}}{1-\tilde{\delta}} + \ln(1-\tilde{\delta}) \right), \\ |\phi'(\mu, \mathbf{x})| &\leq -\sqrt{(1+K)(m_1\varsigma_1 + m_2\varsigma_2)} \ln(1-\tilde{\delta}). \end{aligned}$$

The proof of Theorem 5.2 makes also use of the following lemma, in which its proof is quite similar to that of [15, Lemma 7].

**Lemma 7.4.** *Let  $\mu > 0$  and  $x \in \mathcal{F}_0$  be such that  $\tilde{\delta} < 1$ , where  $\tilde{\delta}$  is as defined in Lemma 7.3. Let  $\mu^+ := \gamma\mu$  with  $\gamma \in (0, 1)$ . Then*

$$\eta(\mu^+, x) - \eta(\mu^+, x(\mu^+)) \leq O((1 + K)(n_1\varsigma_1 + n_2\varsigma_2))\mu^+.$$

Observe that the previous lemma requires  $\tilde{\delta} < 1$ . However, evaluating  $\tilde{\delta}$  explicitly may not be possible. Now we will see that  $\tilde{\delta}$  is actually proportional to  $\delta$ , which can be evaluated. The following lemma is due to [15, Lemma 8].

**Lemma 7.5.** *For any  $\mu > 0$  and  $x \in \mathcal{F}_0$ , we denote  $\tilde{\Delta}x := x - x(\mu)$  and define*

$$\tilde{\delta} := \tilde{\delta}(\mu, x) = \sqrt{\frac{1}{\mu} \nabla_{xx}^2 \eta(\mu, x) [\tilde{\Delta}x, \tilde{\Delta}x]}.$$

*If  $\delta < 1/6$ , then  $\frac{2}{3}\delta \leq \tilde{\delta} \leq 2\delta$ .*

Theorem 5.2 follows directly by combining Lemmas 7.1(ii), 7.4, and 7.5.

## References

- [1] M.A.H. Dempster, *Stochastic Programming*, Academic Press, London, UK, 1980.
- [2] R.J-B. Wets, *Stochastic programming: Solution techniques and approximation schemes*, in: A. Bachem, M. Groeschel, B. Korte (Eds.), *Mathematical Programming: The State of the Art*, Springer-Verlag, Berlin, 1982, 566–603.
- [3] M.A. Goberna, M.A. López, *Linear Semi-Infinite Optimization*, Wiley, 1998.
- [4] R. Hettich, K.O. Kortanek, *Semi-infinite programming: Theory, methods, and applications*, *SIAM Review*, 35 (3) (1993) 380–429.
- [5] S.A. El-Wakil, A.R. Degheidy, M. Sallah, *Time-dependent radiation transfer in a semi-infinite stochastic medium with Rayleigh scattering*, *J. Quant. Spectrosc. Radiat. Transfer*. 85 (1) (2004) 13–24.
- [6] M.M. Selim, M. Sallah, *Neutron transport through semi-infinite continuous stochastic media using Gaussian statistics*, *Ann. Nucl. Energy*. 35 (9) (2008) 1613–1620
- [7] P. Guo, G.H. Huang, L. He, H.L. Li, *Interval-parameter fuzzy-stochastic semi-infinite mixed-integer linear programming for waste management under uncertainty*, *Environ. Model. Assess.* 14 (4) (2008) 521–537.
- [8] Z.-Q. Luo, C. Roos, T. Terlaky, *Complexity analysis of logarithmic barrier decomposition methods for semi-infinite linear programming*, *Appl. Numer. Math.* 29 (1999) 379–394.
- [9] J. Kaliski, D. Haglin, C. Roos, T. Terlaky, *A logarithmic barrier decomposition method for semi-infinite programming*, *Internat. Trans. Oper.Res.* 4 (4) (1997) 285–303.

- [10] D. den Hertog, C. Roos, T. Terlaky, *A build-up variant of the logarithmic barrier method for LP*, Oper. Res. Lett. 12 (1992) 181–186.
- [11] P.M. Vaidya, *A new algorithm for minimizing convex functions over convex sets*, Math. Program. Ser. A 73 (3) (1996) 291–341.
- [12] K.M. Anstreicher, *Volumetric path following algorithms for linear programming*, Math. Program. Ser. A 76 (1997) 245–263.
- [13] A.J. Felt, *A Computational Evaluation of Interior Point Cutting Plane Algorithms for Stochastic Programs*, Ph.D. dissertation, Department of Pure and Applied Mathematics, Washington State University, Pullman, WA, August 2000.
- [14] K.A. Ariyawansa, Y. Zhu, *A class of volumetric barrier decomposition algorithms for stochastic quadratic programming*, Appl. Math. Comput. 186 (2007) 1683–1693.
- [15] K.A. Ariyawansa, Y. Zhu, *A class of polynomial volumetric barrier decomposition algorithms for stochastic semidefinite programming*, Math. of Comp. 80 (2011) 1639–1661.
- [16] B. Alzalg, *Volumetric barrier decomposition algorithms for stochastic quadratic second-order cone programming*, Appl. Math. Comput. 256 (2015) 494–508.
- [17] C. Roos, J.-Ph. Vial, *A polynomial method of approximate centers for linear programming*, Math. Program. 54 (1992) 295–305.
- [18] S.-A. Gustafson, *On numerical analysis in semi-infinite programming*, in: R. Hettich, ed., *Semi-Infinite Programming*, Lecture Notes in Control and Information Sciences, Springer, New York, 15 (1979) 51–65.
- [19] K.M. Anstreicher, *Large step volumetric potential reduction algorithms for linear programming*, Annals of O.R. 62 (1996) 521–538.
- [20] Yu.E. Nesterov, A.S. Nemirovskii, *Interior Point Polynomial Algorithms in Convex Programming*, SIAM Publications, Philadelphia, PA, 1994.