JOTA manuscript No.
(will be inserted by the editor)

Logarithmic-Barrier Decomposition Interior-Point Methods for Stochastic Linear Optimization in a Hilbert Space

Baha Alzalg

Received: 2 April 2019 / Accepted: (will be inserted by the editor)

Abstract Several logarithmic-barrier interior-point methods are now available for solving two-stage stochastic optimization problems with recourse in the finite-dimensional setting. However, despite the genuine need for studying such methods in general spaces, there are no infinite-dimensional analogs of these methods. Inspired by this evident gap in the literature, in this paper, we propose logarithmic-barrier decomposition-based interior-point algorithms for two-stage stochastic linear optimization problems with recourse in a Hilbert space. We study the fundamental properties of the logarithmic barrier associated with the recourse function of our problem setting. The novelty of our algorithms is that their iteration complexity results are independent on the choice of the underlying Hilbert space. In other words, after applying the obtained fundamental properties to our problem setting, the iteration complexity results obtained for the short- and long-step algorithms coincide with the best-known estimates in the finite-dimensional case.

Keywords Linear programming · Stochastic programming · Programming in abstract spaces · Hilbert space · Interior-point methods

Mathematics Subject Classification (2000) 90C05 · 90C15 · 90C48 · 46C07 · 90C51

✉ Baha Alzalg[1,2]
b.alzalg@ju.edu.jo; bmnsma@rit.edu

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

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

1 Introduction

Numerous important models lead to constrained optimization problems in general spaces. Inverse problems, control problems, and shape optimization are only a few prototypical examples. This inspired many authors to study interior-point methods (IPMs, for short) in general spaces. Since Hilbert spaces provide a natural extension of finite-dimensional Euclidean spaces, most of the available results on interior-point methods are given in the setting of Hilbert spaces, see [1–8]. In particular, the work of Renegar [1] (see also [2]) is an infinite-dimensional extension of the interior-point methodology given by Nesterov and Nemirovskii [9] for the finite-dimensional case. Other notable examples are the work of Faybusovich and Moore [3, 4] and Lim and Moore [5] which extend IPMs for optimization problems in a finite-dimensional Euclidean space to a Hilbert space setting.

On the other hand, very active research has been going on in developing numerous efficient decomposition IPMs for solving different two-stage stochastic optimization problems with recourse. In the following, we briefly summarize some of the decomposition algorithmic results. We begin this discussion with two-stage stochastic linear programming and then we will pass to its nonlinear extensions. Zhao [10] derived an IPM for stochastic linear programming, Cho [11] and Ariyawansa and Zhu [12] developed IPMs for stochastic (convex) quadratic linear programming, Alzalg [13, 14] developed IPMs for stochastic second-order cone programming, Mehrotra and Özevin [15] and Ariyawansa and Zhu [16] developed IPMs for stochastic semidefinite programming, Alzalg and Ariyawansa [17] developed an IPM for stochastic symmetric cone programming, Zhao [18] developed an IPM for stochastic convex programming, and finally Alzalg et al. [19] developed an IPM for stochastic linear semi-infinite programming. We emphasize that all the stochastic optimization problems considered in [10–19] and mentioned above are studied in finite-dimensional spaces. To the best of our knowledge, there are no published studies of IPMs for stochastic optimization problems posed in the infinite-dimensional spaces. In this context, we also direct the reader's attention to the unpublished work of Alzalg [20] which studies an IPM for stochastic programming in the particular case of spin factors.

Motivated by the apparent gap related to IPMs for stochastic optimization problem in general spaces, in this paper, we derive logarithmic-barrier decomposition-based interior-point algorithms for two-stage stochastic linear optimization problems with recourse in a Hilbert space setting. Our results can be considered as an extension of the results given by Zhao [10] for stochastic linear optimization problems in a finite-dimensional Euclidean space. Although there is some harmony between the Hilbert space and Euclidean space, our generalization is challenging for the following reasons. Firstly, we deal with an (abstract) inner product instead of the (usual) dot product and deal with operators acting on elements of a Hilbert space instead of matrices acting on vectors. Secondly, the notion of the Fréchet derivative (which is a derivative on Hilbert spaces) is highly involved in our computation instead of the usual derivative. Finally and most notably, we deal with self-concordance barriers defined in an open nonempty convex subspace of an infinite dimensional real Hilbert space instead of those which were defined in an open nonempty convex subset of a finite dimensional real vector space (see Nesterov and Nemirovskii [9, Definition 2.1.1]). The subsequent development will show the technical depth of the self-concordance analysis in this paper.

We note that the corresponding self-concordance analysis of Nesterov and Nemirovskii [9] for the finite-dimensional case can be found in the work of Renegar [1] (see also Renegar [2] and Faybusovich and Moore [4]) for the infinite-dimensional case. In this paper, we study the self-concordance properties of the logarithmic barrier as fundamental properties associated with the recourse function of our problem setting. These properties will play a crucial role in developing polynomial-time interior-point algorithms for our problem. We also consider short- and long-step variants of path-following algorithms and obtain estimates on the number of Newton iterations needed to find an approximate of the optimal solution. The iteration complexity results obtained for the short- and long-step algorithms coincide with the best known in the finite-dimensional case obtained in Zhao [10]. In fact, we will see that the complexity results of the proposed algorithms are independent of the choice of the underlying Hilbert space. This is the main contribution and novelty of our methods.

The organization of this paper is as follows. In Section 2, we present the necessary background material and notations. Section 3 presents problem formulation and introduces some assumptions. We also describe in Section 3 the time-dependent linear stochastic control problem as an example model leading to our stochastic optimization problem. In Section 4, we compute the Fréchet derivatives of the recourse function. Section 5 gives some fundamental properties of the recourse function. In Section 6, we formally state the proposed algorithms and present the complexity results for the short- and long-step versions of the algorithms. Section 7 contains some concluding remarks.

2 Preliminaries

Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a real Hilbert space with identity operator $\mathcal{I}$, and let $\mathcal{G}$ be a closed subspace of $\mathcal{H}$. The *orthogonal complement* of $\mathcal{G}$ is defined by

$$\mathcal{G}^{\perp} := \{\boldsymbol{x} \in \mathcal{H} : \boldsymbol{x} \perp \mathcal{G}\}$$

and contains all elements in $\mathcal{H}$ that are orthogonal to every element in $\mathcal{G}$. It is known that $(\mathcal{G}^{\perp})^{\perp} = \mathcal{G}$, $\mathcal{G} \cap \mathcal{G}^{\perp} = \{\mathbf{0}\}$ and that $\mathcal{H} = \mathcal{G} \oplus \mathcal{G}^{\perp}$.

The inner product $\langle\cdot,\cdot\rangle$ on $\mathcal{H}$ defines a *norm* on $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ given by

$$||\boldsymbol{x}|| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}, \text{ for } \boldsymbol{x} \in \mathcal{H}.$$

The following theorem is of fundamental importance for our subsequent development.

Theorem 2.1 [21, Theorem 3.3-1] *Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a real Hilbert space and let $\mathcal{G}$ be a nonempty subspace of $\mathcal{H}$. Then for every given (fixed) element $\boldsymbol{x} \in \mathcal{H}$ there exists a unique element $\bar{\boldsymbol{y}} \in \mathcal{G}$ such that*

$$\inf_{\boldsymbol{y} \in \mathcal{G}} ||\boldsymbol{x} - \boldsymbol{y}|| = ||\boldsymbol{x} - \bar{\boldsymbol{y}}||. \tag{1}$$

Moreover, the element $\boldsymbol{x} - \bar{\boldsymbol{y}} \in \mathcal{G}^{\perp}$.

Let $\boldsymbol{x}, \boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m \in \mathcal{H}$. Throughout this paper, we denote by "$\mathcal{A}\boldsymbol{x}$" the vector in $\mathbb{R}^m$ whose i^{th}-entry is the scalar $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle, i = 1, 2, \ldots, m$. For $\boldsymbol{z} \in \mathbb{R}^m$, we also denote by "$\mathcal{A}^{\dagger}\boldsymbol{z}$" the element $\sum_{i=1}^{m} z_i \boldsymbol{a}_i$ in $\mathcal{H}$. Note that

$$\left\langle \mathcal{A}^{\dagger}\boldsymbol{z}, \boldsymbol{x} \right\rangle = \left\langle \sum_{i=1}^{m} z_i \boldsymbol{a}_i, \boldsymbol{x} \right\rangle = \sum_{i=1}^{m} z_i \langle \boldsymbol{a}_i, \boldsymbol{x}\rangle = \boldsymbol{z}^{\mathsf{T}} \mathcal{A}\boldsymbol{x}. \tag{2}$$

Note also that $\mathcal{A}$ maps the Hilbert space $\mathcal{H}$ onto $\mathbb{R}^m$, while $\mathcal{A}^{\dagger}$ maps $\mathbb{R}^m$ into the Hilbert space $\mathcal{H}$.

An operator $\mathcal{T} : \mathcal{H} \longrightarrow \mathcal{H}$ is said to be *invertible* if there exists another operator, called the *inverse operator* of $\mathcal{T}$ and denoted by $\mathcal{T}^{-1} : \mathcal{H} \longrightarrow \mathcal{H}$, such that $\mathcal{T}\mathcal{T}^{-1} = \mathcal{T}^{-1}\mathcal{T} = \mathcal{I}$.

Throughout the paper we use "$\mathcal{D}_x$", "$\mathcal{D}^2_{xx}$"and "$\mathcal{D}^3_{xxx}$" to denote the first, second and third Fréchet derivatives with respect to x, respectively. The above derivatives are defined in the standard way. We have the following remark (see also [4, Section 2]).

Remark 2.1 Let $(\mathcal{H}, \langle\cdot,\cdot\rangle)$ be a Hilbert space and $x \in \mathcal{H}$. Let also $f : \mathcal{H} \longrightarrow \mathbb{R}$ be a smooth function, then the gradient $\nabla_x f(x)$ is uniquely determined by

$$\mathcal{D}_x f(x)(\xi) := \langle \nabla_x f(x), \xi \rangle$$

for any $\xi \in \mathcal{H}$. Here, $\mathcal{D}_x f(x)(\xi)$ stands for the first Fréchet derivative of f at the element x evaluated on ξ. The second Fréchet derivative is the derivative of the first Fréchet derivative, which is given by

$$\mathcal{D}^2_{xx} f(x)(\xi, \zeta) := \mathcal{D}_x \left(\mathcal{D}_x f(x)(\xi) \right)(\zeta),$$

for any $\xi, \zeta \in \mathcal{H}$. Here, $\mathcal{D}^2_{xx} f(x)(\xi, \zeta)$ stands for the second Fréchet derivative of f at the element x evaluated on (ξ, ζ). Higher Fréchet derivatives are defined in a similar way.

In particular, we have

$$\mathcal{D}_x \left\langle \mathcal{A}^\dagger z, x \right\rangle(\xi) = \mathcal{D}_x \left(z^\mathsf{T} \mathcal{A} x \right)(\xi) = \mathcal{D}_x \left(\sum_{i=1}^m \langle a_i, x \rangle z_i \right)(\xi) = \sum_{i=1}^m \langle a_i, \xi \rangle z_i = \left\langle \sum_{i=1}^m z_i a_i, \xi \right\rangle = \left\langle \mathcal{A}^\dagger z, \xi \right\rangle,$$

and hence

$$\mathcal{D}_x \left\langle \mathcal{A}^\dagger z, x \right\rangle = \mathcal{D}_x \left(z^\mathsf{T} \mathcal{A} x \right) = \mathcal{A}^\dagger z. \tag{3}$$

If w is an m^{th}-dimensional vector whose each component is a real functional defined on a Hilbert space $\mathcal{H}$ (i.e., $w = w(x) \in \mathbb{R}^m$ for $x \in \mathcal{H}$), then the Fréchet derivative $\mathcal{D}_x w_i$ and gradient $\nabla_x w_i$ reside the Hilbert space $\mathcal{H}$. For $z \in \mathbb{R}^m$, we denote by $\mathcal{D}_x w^\dagger \circ z$ the element $\sum_{i=1}^m z_i \mathcal{D}_x w_i$ in $\mathcal{H}$. For example

$$\mathcal{D}_x \left(z^\mathsf{T} w \right) = \mathcal{D}_x \sum_{i=1}^m z_i w_i = \sum_{i=1}^m z_i \mathcal{D}_x w_i = \mathcal{D}_x w^\dagger \circ z.$$

In particular, we generalize (3) by replacing $\boldsymbol{x} \in \mathcal{H}$ with $\boldsymbol{y} = \boldsymbol{y}(\boldsymbol{x}) \in \mathcal{H}$ and get

$$\mathcal{D}_{\boldsymbol{x}} \left\langle \mathcal{A}^{\dagger} \boldsymbol{z}, \boldsymbol{y} \right\rangle = \mathcal{D}_{\boldsymbol{x}} \left(\boldsymbol{z}^{\mathsf{T}} \mathcal{A} \boldsymbol{y} \right) = \mathcal{D}_{\boldsymbol{x}} (\mathcal{A} \boldsymbol{y})^{\dagger} \circ \boldsymbol{z}. \tag{4}$$

We write $\boldsymbol{e}$ for a vector with all entries equal to one (its dimension will be clear from the context). For $\boldsymbol{\xi} \in \mathcal{H}$ and $\boldsymbol{w} = \boldsymbol{w}(\boldsymbol{x}) \in \mathbb{R}^m$ with $\boldsymbol{x} \in \mathcal{H}$, we denote by $\mathcal{D}_{\boldsymbol{x}} \boldsymbol{w} \otimes \boldsymbol{\xi}$ the vector in $\mathbb{R}^m$ whose i^{th}-entry is the scalar $\langle \nabla_{\boldsymbol{x}} w_i, \boldsymbol{\xi} \rangle$. For example

$$\mathcal{D}_{\boldsymbol{x}} \left(\boldsymbol{e}^{\mathsf{T}} \boldsymbol{w} \right) (\boldsymbol{\xi}) = \left\langle \nabla_{\boldsymbol{x}} \boldsymbol{e}^{\mathsf{T}} \boldsymbol{w}, \boldsymbol{\xi} \right\rangle = \left\langle \nabla_{\boldsymbol{x}} \left(\sum_{i=1}^{m} w_i \right), \boldsymbol{\xi} \right\rangle = \sum_{i=1}^{m} \langle \nabla_{\boldsymbol{x}} w_i, \boldsymbol{\xi} \rangle = \boldsymbol{e}^{\mathsf{T}} \left(\mathcal{D}_{\boldsymbol{x}} \boldsymbol{w} \otimes \boldsymbol{\xi} \right).$$

For any strictly positive vector $\boldsymbol{x} \in \mathbb{R}^m$, we define $\ln \boldsymbol{x} := (\ln x_1, \ln x_2, \ldots, \ln x_m)^{\mathsf{T}}$ and set $\boldsymbol{x}^{-1} := (x_1^{-1}, x_2^{-1}, \ldots, x_m^{-1})^{\mathsf{T}}$. We let $X := \operatorname{diag}(x_1, x_2, \ldots, x_m)$ denote the $m \times m$ diagonal matrix whose diagonal entries are $x_1, x_2, \ldots, x_m$.

We will also use the following notations to compare of our results with the results from [11, 13, 15, 17]:

- $\mathbb{R}^m_+$ denotes the nonnegative orthant cone of $\mathbb{R}^m$.
- $\mathcal{Q}^m$ denotes the second-order cone [13] of dimension m.
- $\mathcal{S}^m$ denotes the cone of symmetric positive semidefinite matrices of order m.
- $\mathcal{K}^r$ denotes a symmetric cone [17] of rank r.
- G^{ς} denotes a convex set admitting a non-degenerate and strongly self-concordance barrier function with a complexity value ς.

It is known that G^{ς} includes $\mathcal{K}^r$ as a special case, $\mathcal{K}^r$ includes $\mathcal{S}^m$ as a special case, $\mathcal{S}^m$ includes $\mathcal{Q}^m$ as a special case, and $\mathcal{Q}^m$ includes $\mathbb{R}^m_+$ as a special case.

3 Problem Formulation and Assumptions

In this section, we introduce the two-stage stochastic linear optimization problem with recourse in a Hilbert space. We also describe the time-dependent linear stochastic control problem as an example model leading to our optimization problem. Then we present the logarithmic-barrier problem associated with our formulation, define some feasibility sets and make some assumptions.

3.1 Problem Formulation

Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a real Hilbert space with identity operator $\mathcal{I}$. Let $\mathcal{G}$ be a (deterministic) closed subspace $\mathcal{H}$ and $\mathcal{G}(\omega)$ be a random closed subspace of $\mathcal{H}$ whose realizations depend on an underlying outcome ω in an event space Ω with a known probability function P. For $1 \leq i \leq m_1$ and $1 \leq j \leq m_2$, we also let $\boldsymbol{c}, \boldsymbol{a}_i, \boldsymbol{t}_j \in \mathcal{H}$ and $b_i \in \mathbb{R}$ be deterministic data, and $\boldsymbol{d}, \boldsymbol{w}_j \in \mathcal{H}$ and $h_j \in \mathbb{R}$ be random data whose realizations depend on an underlying outcome ω in an event space Ω with a known probability function P. We consider the *two-stage stochastic linear optimization problem with recourse over the Hilbert space* $\mathcal{H}$, which has the form

$$\begin{aligned} &\min\ \langle \boldsymbol{c}, \boldsymbol{x} \rangle + \mathbb{E}\left[\varrho(\boldsymbol{x}, \omega)\right] \\ &\text{s.t.}\ \ \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \leq b_i,\ i = 1, 2, \ldots, m_1, \\ &\qquad \boldsymbol{x} \in \mathcal{G}, \end{aligned} \tag{5}$$

where $\boldsymbol{x}$ is the first-stage decision variable, and $\varrho(\boldsymbol{x}, \omega)$ is the minimum value of the problem

$$\begin{aligned} &\min\ \langle \boldsymbol{d}(\omega), \boldsymbol{y}(\omega) \rangle \\ &\text{s.t.}\ \ \langle \boldsymbol{w}_j(\omega), \boldsymbol{y}(\omega) \rangle \leq h_j(\omega) - \langle \boldsymbol{t}_j(\omega), \boldsymbol{x} \rangle,\ j = 1, 2, \ldots, m_2, \\ &\qquad \boldsymbol{y}(\omega) \in \mathcal{G}(\omega), \end{aligned} \tag{6}$$

where $\boldsymbol{y}(\omega)$ is the second-stage variable, and

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

We begin by writing the extensive formulation of Problem (5, 6). We examine Problem (5, 6) when the event space Ω is finite. Let $\mathcal{G}^{(1)}, \mathcal{G}^{(2)}, \ldots, \mathcal{G}^{(K)}$ be K possible realizations of the random closed subspace $\mathcal{G}(\omega)$ and let $p_k := P\left(\mathcal{G}(\omega) = \mathcal{G}^{(k)}\right)$ be the associated probability for $k = 1, 2, \ldots, K$. Similarly, let $\left\{\left(\boldsymbol{t}_j^{(k)}, \boldsymbol{w}_j^{(k)}, h_j^{(k)}, \boldsymbol{d}^{(k)}\right) : k = 1, 2, \ldots, K\right\}$ be the set of K possible realizations of the random variables $\big(\boldsymbol{t}_j(\omega), \boldsymbol{w}_j(\omega),$ $h_j(\omega), \boldsymbol{d}(\omega)\big)$ and let $p_k := P\left(\left(\boldsymbol{t}_j(\omega), \boldsymbol{w}_j(\omega), h_j(\omega), \boldsymbol{d}(\omega)\right) = \left(\boldsymbol{t}_j^{(k)}, \boldsymbol{w}_j^{(k)}, h_j^{(k)}, \boldsymbol{d}^{(k)}\right)\right)$ be the associated probability for

$k = 1, 2, \ldots, K$. Then Problem (5, 6) becomes

$$\begin{aligned} \min\ & \eta(\boldsymbol{x}) := \langle \boldsymbol{c}, \boldsymbol{x} \rangle + \sum_{k=1}^{K} \varrho^{(k)}(\boldsymbol{x}) \\ \text{s.t.}\ & \langle \boldsymbol{a}_i, \boldsymbol{x} \rangle \leq b_i,\ i = 1, 2, \ldots, m_1, \\ & \boldsymbol{x} \in \mathcal{G}, \end{aligned} \tag{7}$$

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

$$\begin{aligned} \min\ & \left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)} \right\rangle \\ \text{s.t.}\ & \left\langle \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)} \right\rangle \leq h_j^{(k)} - \left\langle \boldsymbol{t}_j^{(k)}, \boldsymbol{x} \right\rangle,\ j = 1, 2, \ldots, m_2, \\ & \boldsymbol{y}^{(k)} \in \mathcal{G}^{(k)}, \end{aligned} \tag{8}$$

where p_k is absorbed by $\varrho^{(k)}(\boldsymbol{x})$ in (7), which is possible by redefining $\boldsymbol{d}^{(k)}$ as $\boldsymbol{d}^{(k)} := p_k \boldsymbol{d}^{(k)}$ in (8), for each $k = 1, 2, \ldots, K$.

The dual of Problem (8) is the problem

$$\begin{aligned} \max\ & \sum_{j=1}^{m_2} \left(\left\langle \boldsymbol{t}_j^{(k)}, \boldsymbol{x} \right\rangle - h_j^{(k)} \right) z_j^{(k)} \\ \text{s.t.}\ & \boldsymbol{d}^{(k)} + \sum_{j=1}^{m_2} z_j^{(k)} \boldsymbol{w}_j^{(k)} \in \mathcal{G}^{(k)\perp}, \\ & \boldsymbol{z}^{(k)} \geq \boldsymbol{0}, \end{aligned} \tag{9}$$

where $\boldsymbol{z}^{(k)} \in \mathbb{R}^{m_2}$ is the second-stage dual multiplier.

3.2 An Example Model

The application example presented in this part can be viewed as the linear stochastic version of the time-dependent linear deterministic control problem described by Faybusovich and Moore [3, Section 4]. Han and Van Roy [22] and Ahn and Haugh [23] proposed linear programming approaches for the approximate solution of the Hamilton-Jacobi-Bellman equation that arises from continuous-time control problems. Their approaches apply to diffusion problems with a horizon T. The horizon, T, is assumed to be finite, but it may be either deterministic or exponentially distributed random variable [22, 23]. Therefore, it is meaningful to describe a linear stochastic version of the time-dependent linear deterministic control problem described in [3, Section 4].

Let $\mathcal{H} := L_2^n(\mathbb{R})$ be the Hilbert space of measurable square integrable functions on $\mathbb{R}$ with values in $\mathbb{R}^n$. This Hilbert space is equipped with the inner product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \int_S \boldsymbol{x}(t)^\top \boldsymbol{y}(t)\, dt$$

for $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{H}$. Here, $S \subseteq \mathbb{R}$ is a subset over which we integrate. Let also

$$\begin{aligned} \mathcal{G} \quad &:= \{\boldsymbol{x} \in \mathcal{H} : \boldsymbol{x} \text{ is absolutely continuous on } [0, T]\}, \\ \mathcal{G}^{(k)} &:= \left\{\boldsymbol{y} \in \mathcal{H} : \boldsymbol{y} \text{ is absolutely continuous on } \left[0, T^{(k)}\right]\right\}, k = 1, 2, \ldots, K, \end{aligned}$$

where $T, T^{(1)}, \ldots, T^{(K)}$ are fixed positive numbers.

Depending on the input data from the modeler of the application components, it is also possible to include initial and boundary conditions in the definitions of the closed subspaces $\mathcal{G}$ and $\mathcal{G}^{(k)}$, such as $\boldsymbol{x}(0) = \alpha$ and $\boldsymbol{y}(0) = \beta$ for $\alpha, \beta \in \mathbb{R}$.

For $1 \leq i \leq m_1, 1 \leq j \leq m_2$ and $1 \leq k \leq K$, we let $\boldsymbol{c}, \boldsymbol{a}_i, \boldsymbol{t}_j^{(k)}, \boldsymbol{d}^{(k)}, \boldsymbol{w}_j^{(k)} \in \mathcal{H}$ and $b_i, h_j^{(k)} \in \mathbb{R}$ be given data. The vectors-valued functions $\boldsymbol{c}, \boldsymbol{a}_i$ and $\boldsymbol{t}_j^{(k)}$ are continuous on $[0, T]$ with values $\mathbb{R}^n$, and the vectors-valued functions $\boldsymbol{d}^{(k)}$ and $\boldsymbol{w}_j^{(k)}$ are continuous on $\left[0, T^{(K)}\right]$ with values $\mathbb{R}^n$.

Given this, a stochastic two-stage time-dependent linear control problem with K scenarios has the form

$$\begin{aligned} \min \quad & \int_0^T \boldsymbol{c}(t)^\top \boldsymbol{x}(t)\, dt + \sum_{k=1}^K \varrho^{(k)}(\boldsymbol{x}) \\ \text{s.t.} \quad & \int_0^T \boldsymbol{a}_i(t)^\top \boldsymbol{x}(t)\, dt \leq b_i,\ i = 1, 2, \ldots, m_1, \\ & \boldsymbol{x} \in \mathcal{G}, \end{aligned} \tag{10}$$

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

$$\begin{aligned} \min\; & \int_0^{T^{(k)}} \boldsymbol{d}^{(k)}(t)^{\mathsf{T}} \boldsymbol{y}^{(k)}(t)\, dt \\ \text{s.t.}\; & \int_0^{T^{(k)}} \boldsymbol{w}_j^{(k)}(t)^{\mathsf{T}} \boldsymbol{y}^{(k)}(t)\, dt \le h_j^{(k)} - \int_0^{T} \boldsymbol{t}_j^{(k)}(t)^{\mathsf{T}} \boldsymbol{x}(t)\, dt,\; j = 1, 2, \ldots, m_2, \\ & \boldsymbol{y}^{(k)} \in \mathcal{G}^{(k)}. \end{aligned} \tag{11}$$

This example is generic in the sense that it applies to high-dimensional control problems in different diffusion-based settings, and covers portfolio optimization problems where we choose a self-financing trading strategy that maximizes the expected utility of terminal wealth. See, for instance, the linear optimization models presented in [22, 23], and the references contained therein.

3.3 The Logarithmic-Barrier Problem

Following our notations in Section 2, we denote by $\mathcal{A}\boldsymbol{x}$ the vector in $\mathbb{R}^{m_1}$ whose i^{th}-entry is $\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle, i = 1, 2, \ldots, m_1$, and denote by $\mathcal{T}^{(k)}\boldsymbol{x}$ and $\mathcal{W}^{(k)}\boldsymbol{y}^{(k)}$ the vectors in $\mathbb{R}^{m_2}$ whose j^{th}-entries are $\left\langle \boldsymbol{t}_j^{(k)}, \boldsymbol{x}\right\rangle$ and $\left\langle \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)}\right\rangle$, respectively, for $j = 1, 2, \ldots, m_2$ and $k = 1, 2, \ldots, K$.

The logarithmic-barrier problem associated with Problem (7, 8) is

$$\begin{aligned} \min\; & \eta(\mu, \boldsymbol{x}) := \langle \boldsymbol{c}, \boldsymbol{x}\rangle + \sum_{k=1}^{K} \rho^{(k)}(\mu, \boldsymbol{x}) - \mu\, \boldsymbol{e}^{\mathsf{T}} \ln \boldsymbol{s} \\ \text{s.t.}\; & \mathcal{A}\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b}, \\ & \boldsymbol{x} \in \mathcal{G},\; \boldsymbol{s} > \boldsymbol{0}, \end{aligned} \tag{12}$$

where $\mu > 0$ is a barrier parameter, and, for $k = 1, 2, \ldots, K$, $\rho^{(k)}(\mu, \boldsymbol{x})$ is the minimum value of the problem

$$\begin{aligned} \min\; & \left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle - \mu\, \boldsymbol{e}^{\mathsf{T}} \ln \boldsymbol{s}^{(k)} \\ \text{s.t.}\; & \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} + \boldsymbol{s}^{(k)} = \boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x}, \\ & \boldsymbol{y}^{(k)} \in \mathcal{G}^{(k)},\; \boldsymbol{s}^{(k)} > \boldsymbol{0}. \end{aligned} \tag{13}$$

Here, $\boldsymbol{s}$ and $\boldsymbol{s}^{(k)}$ are the first- and second-stage slack decision variables, respectively. If for some k, Problem (13) is infeasible, then we define $\sum_{k=1}^{K} \rho^{(k)}(\mu, \boldsymbol{x}) := \infty$.

Following our notations in Section 2, we also denote by $\mathcal{W}^{(k)\dagger}\boldsymbol{z}^{(k)}$ the element $\sum_{j=1}^{m_2} z_j^{(k)}\boldsymbol{w}_j^{(k)}$ in the subspace $\mathcal{G}^{(k)}$ of the Hilbert space $\mathcal{H}$.

The Lagrangian dual of Problem (13) is the problem

$$
\begin{aligned}
\max\ & \left(\mathcal{T}^{(k)}\boldsymbol{x} - \boldsymbol{h}^{(k)}\right)^{\mathsf{T}} \boldsymbol{z}^{(k)} + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{z}^{(k)} \\
\text{s.t.}\ \ & \boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger}\boldsymbol{z}^{(k)} \in \mathcal{G}^{(k)\perp}, \\
& \boldsymbol{z}^{(k)} > \boldsymbol{0},
\end{aligned}
\tag{14}
$$

which is the logarithmic-barrier problem associated with Problem (9). Here $\boldsymbol{z}^{(k)} \in \mathbb{R}^{m_2}$ is the second-stage dual multiplier.

Note that $\left(\boldsymbol{y}^{(k)}, \boldsymbol{s}^{(k)}\right)$ and $\boldsymbol{z}^{(k)}$ are optimal solutions to (13) and (14), respectively, iff they satisfy the following optimality conditions:

$$
\begin{aligned}
& \mathcal{S}^{(k)}\boldsymbol{z}^{(k)} = \mu\, \boldsymbol{e}, \\
& \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} + \boldsymbol{s}^{(k)} = \boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x}, \\
& \left\langle \boldsymbol{y}^{(k)},\ \boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger}\boldsymbol{z}^{(k)} \right\rangle = 0, \\
& \boldsymbol{s}^{(k)} > \boldsymbol{0},\ \boldsymbol{z}^{(k)} > \boldsymbol{0},
\end{aligned}
\tag{15}
$$

where the third equality follows from the fact that the variable $\boldsymbol{y}^{(k)}$ belongs to $\mathcal{G}^{(k)}$ and the element $\boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger}\boldsymbol{z}^{(k)}$ belongs to its orthogonal complement $\mathcal{G}^{(k)\perp}$.

Problem (12, 13) can be equivalently written as a deterministic optimization problem

$$
\begin{aligned}
\min\ & \eta(\mu, \boldsymbol{x}) = \langle \boldsymbol{c}, \boldsymbol{x}\rangle - \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{s} + \sum_{k=1}^{K} \overbrace{\left(\left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle - \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{s}^{(k)}\right)}^{\rho^{(k)}(\mu,\boldsymbol{x})} \\
\text{s.t.}\ \ & \mathcal{A}\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b}, \\
& \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} + \boldsymbol{s}^{(k)} = \boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x},\ k = 1, 2, \ldots, K, \\
& \boldsymbol{x} \in \mathcal{G},\ \boldsymbol{s} > \boldsymbol{0}, \\
& \boldsymbol{y}^{(k)} \in \mathcal{G}^{(k)},\ \boldsymbol{s}^{(k)} > \boldsymbol{0},\ k = 1, 2, \ldots, K.
\end{aligned}
\tag{16}
$$

We now define the following feasibility sets

$$\begin{aligned}
\mathcal{F}_1 &:= \left\{ \boldsymbol{x} \in \mathcal{G} : \boldsymbol{s} = \boldsymbol{b} - \mathcal{A}\boldsymbol{x} > \boldsymbol{0} \right\};\\
\mathcal{F}_2^{(k)}(\boldsymbol{x}) &:= \left\{ \boldsymbol{y}^{(k)} \in \mathcal{G}^{(k)} : \boldsymbol{s}^{(k)} = \boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x} - \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} > \boldsymbol{0} \right\} \text{ for } k = 1, 2, \ldots, K;\\
\mathcal{F}_2^{(k)} &:= \left\{ \boldsymbol{x} \in \mathcal{G} : \mathcal{F}_2^{(k)}(\boldsymbol{x}) \neq \emptyset \right\} \text{ for } k = 1, 2, \ldots, K;\\
\mathcal{F}_2 &:= \textstyle\bigcap_{k=1}^{K} \mathcal{F}_2^{(k)};\\
\mathcal{F}_0 &:= \mathcal{F}_1 \bigcap \mathcal{F}_2;\\
\mathcal{F} &:= \Big\{ (\boldsymbol{x}, \boldsymbol{\lambda}) \times \left(\boldsymbol{y}^{(1)}, \boldsymbol{z}^{(1)}, \boldsymbol{y}^{(2)}, \boldsymbol{z}^{(2)}, \ldots, \boldsymbol{y}^{(K)}, \boldsymbol{z}^{(K)}\right) : \boldsymbol{b} - \mathcal{A}\boldsymbol{x} > \boldsymbol{0}, \boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x} - \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} > \boldsymbol{0},\\
&\qquad \boldsymbol{\lambda} > \boldsymbol{0}, \boldsymbol{z}^{(k)} > \boldsymbol{0}, \left\langle \boldsymbol{y}^{(k)},\ \boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger}\boldsymbol{z}^{(k)} \right\rangle = 0, \left\langle \boldsymbol{x},\ \boldsymbol{c} + \mathcal{A}^{\dagger}\boldsymbol{\lambda} + \textstyle\sum_{k=1}^{K} \mathcal{T}^{(k)\dagger}\boldsymbol{z}^{(k)} \right\rangle = 0,\ k = 1, \ldots, K \Big\}.
\end{aligned}$$

Here $\boldsymbol{\lambda} \in \mathbb{R}^{m_1}$ is the first-stage dual multiplier. Then we make the following two assumptions.

Assumption 3.1 *The elements $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{m_1}$ are linearly independent in $\mathcal{H}$, and for $k = 1, 2, \ldots, K$, the elements $\boldsymbol{t}_1^{(k)}, \boldsymbol{t}_2^{(k)}, \ldots, \boldsymbol{t}_{m_2}^{(k)}$ are linearly independent in $\mathcal{H}$, and the elements $\boldsymbol{w}_1^{(k)}, \boldsymbol{w}_2^{(k)}, \ldots, \boldsymbol{w}_{m_2}^{(k)}$ are linearly independent in $\mathcal{H}$.*

Assumption 3.2 *The feasibility set $\mathcal{F}$ is nonempty.*

Assumption 3.1 is of fundamental importance to ensure invertibility of operators. Assumption 3.2 guarantees strong duality for first- and second-stage stochastic problems by requiring that Problem (16) and its dual have strictly feasible solutions. This means that Problems (12-16) have unique solutions. Note that for a given $\mu > 0$, $\sum_{k=1}^{K} \rho^{(k)}(\mu, \boldsymbol{x}) < \infty$ iff $\boldsymbol{x} \in \mathcal{F}_2$. The optimal solutions of Problems (12, 13) and those of Problem (16) have a relationship given in the following remark. The point $\left(\boldsymbol{x}(\mu); \boldsymbol{y}^{(1)}(\mu), \boldsymbol{y}^{(2)}(\mu), \cdots, \boldsymbol{y}^{(K)}(\mu)\right)$ is the optimal solution of (16) iff $(\boldsymbol{x}(\mu))$ is the optimal solution of (12) and $\left(\boldsymbol{y}^{(1)}(\mu), \boldsymbol{y}^{(2)}(\mu), \cdots, \boldsymbol{y}^{(K)}(\mu)\right)$ are the optimal solutions for (13) for given μ and $\boldsymbol{x} = \boldsymbol{x}(\mu)$.

In the next section, we shall compute the first-order and the second-order Fréchet derivatives of the recourse function $\eta(\cdot, \cdot)$ which will be used to determine the Newton direction and the measure of proximity of the current point $\boldsymbol{x}$ to the central path for our algorithms.

4 Computation of the Fréchet Derivatives of the Recourse Function

In this section, we compute $\mathcal{D}_x\eta(\mu, \boldsymbol{x})$ and $\mathcal{D}^2_{xx}\eta(\mu, \boldsymbol{x})$. To do so, we need to determine the Fréchet derivatives of $\rho^{(k)}(\mu, \boldsymbol{x})$ with respect to $\boldsymbol{x}$. Let $\left(\boldsymbol{y}^{(k)}, \boldsymbol{s}^{(k)}, \boldsymbol{z}^{(k)}\right) := \left(\boldsymbol{y}^{(k)}(\mu, \boldsymbol{x}), \boldsymbol{s}^{(k)}(\mu, \boldsymbol{x}), \boldsymbol{z}^{(k)}(\mu, \boldsymbol{x})\right)$. Using (15), we have that

$$
\begin{aligned}
\left(\boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x}\right)^{\mathsf{T}} \boldsymbol{z}^{(k)} &= \left(\mathcal{W}^{(k)}\boldsymbol{y}^{(k)} + \boldsymbol{s}^{(k)}\right)^{\mathsf{T}} \boldsymbol{z}^{(k)} \\
&= \sum_{j=1}^{m_2} \left\langle z_j^{(k)}\boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)}\right\rangle + \boldsymbol{s}^{(k)^{\mathsf{T}}} \boldsymbol{z}^{(k)} \\
&= \left\langle \boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger}\boldsymbol{z}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle - \left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle + \boldsymbol{s}^{(k)^{\mathsf{T}}} \boldsymbol{z}^{(k)} \\
&= \boldsymbol{s}^{(k)^{\mathsf{T}}} \boldsymbol{z}^{(k)} - \left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle \\
&= \mu m_2 - \left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle.
\end{aligned}
$$

From the definition of $\rho^{(k)}(\cdot, \cdot)$ in (16), it immediately follows that

$$
\begin{aligned}
\left(\mathcal{T}^{(k)}\boldsymbol{x} - \boldsymbol{h}^{(k)}\right)^{\mathsf{T}} \boldsymbol{z}^{(k)} + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{z}^{(k)} &= \left\langle \boldsymbol{d}^{(k)}, \boldsymbol{y}^{(k)}\right\rangle - \mu m_2 + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{z}^{(k)} \\
&= \rho^{(k)}(\mu, \boldsymbol{x}) + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{s}^{(k)} - \mu m_2 + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{z}^{(k)} \\
&= \rho^{(k)}(\mu, \boldsymbol{x}) - \mu m_2 + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln\left(S^{(k)}\boldsymbol{z}^{(k)}\right) \\
&= \rho^{(k)}(\mu, \boldsymbol{x}) - \mu m_2 + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln\left(\mu\boldsymbol{e}\right) \\
&= \rho^{(k)}(\mu, \boldsymbol{x}) - \mu m_2(1 - \ln \mu).
\end{aligned}
$$

Thus,

$$
\rho^{(k)}(\mu, \boldsymbol{x}) = \left(\mathcal{T}^{(k)}\boldsymbol{x} - \boldsymbol{h}^{(k)}\right)^{\mathsf{T}} \boldsymbol{z}^{(k)} + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{z}^{(k)} + \mu m_2(1 - \ln \mu).
$$

It follows that

$$
\begin{aligned}
\mathcal{D}_x\rho^{(k)}(\mu, \boldsymbol{x}) &= \mathcal{D}_x\left(\left(\mathcal{T}^{(k)}\boldsymbol{x} - \boldsymbol{h}^{(k)}\right)^{\mathsf{T}} \boldsymbol{z}^{(k)} + \mu\, \boldsymbol{e}^{\mathsf{T}}\ln \boldsymbol{z}^{(k)}\right) \\
&= \mathcal{T}^{(k)\dagger}\boldsymbol{z}^{(k)} + \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \left(\mathcal{T}^{(k)}\boldsymbol{x} - \boldsymbol{h}^{(k)}\right) + \mu\, \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \boldsymbol{z}^{(k)^{-1}} \\
&= \mathcal{T}^{(k)\dagger}\boldsymbol{z}^{(k)} - \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \left(\mathcal{W}^{(k)}\boldsymbol{y}^{(k)} + \boldsymbol{s}^{(k)}\right) + \mu\, \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \boldsymbol{z}^{(k)^{-1}} \\
&= \mathcal{T}^{(k)\dagger}\boldsymbol{z}^{(k)} - \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} + \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \left(\mu\, \boldsymbol{z}^{(k)^{-1}} - \boldsymbol{s}^{(k)}\right) \\
&= \mathcal{T}^{(k)\dagger}\boldsymbol{z}^{(k)} - \mathcal{D}_x\boldsymbol{z}^{(k)\dagger} \circ \mathcal{W}^{(k)}\boldsymbol{y}^{(k)}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\mathcal{D}_x \boldsymbol{z}^{(k)\dagger} \circ \mathcal{W}^{(k)} \boldsymbol{y}^{(k)} &= \sum_{j=1}^{m_2} \left\langle \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)} \right\rangle \mathcal{D}_x z_j^{(k)} \\
&= \mathcal{D}_x \left(\sum_{j=1}^{m_2} \left\langle \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)} \right\rangle z_j^{(k)} \right) - \sum_{j=1}^{m_2} \left(\mathcal{D}_x \left\langle \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)} \right\rangle \right) z_j^{(k)} \\
&= \mathcal{D}_x \left(\left\langle \sum_{j=1}^{m_2} z_j^{(k)} \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)} \right\rangle \right) - \sum_{j=1}^{m_2} \left(\mathcal{D}_x \left\langle \boldsymbol{w}_j^{(k)}, \boldsymbol{y}^{(k)} \right\rangle \right) z_j^{(k)} \\
&= \mathcal{D}_x \left\langle \mathcal{W}^{(k)\dagger} \boldsymbol{z}^{(k)}, \boldsymbol{y}^{(k)} \right\rangle - \mathcal{D}_x \left(\mathcal{W}^{(k)} \boldsymbol{y}^{(k)} \right)^{\dagger} \circ \boldsymbol{z}^{(k)} \\
&= \mathbf{0},
\end{aligned}
$$

where we used (4). Thus, the first Fréchet derivative of $\rho^{(k)}(\cdot,\cdot)$ reads $\mathcal{D}_x \rho^{(k)}(\mu, \boldsymbol{x}) = \mathcal{T}^{(k)\dagger} \boldsymbol{z}^{(k)}$, and hence

$$
\left\langle \nabla_x \rho^{(k)}(\mu, \boldsymbol{x}), \boldsymbol{\xi} \right\rangle = \left\langle \mathcal{T}^{(k)\dagger} \boldsymbol{z}^{(k)}, \boldsymbol{\xi} \right\rangle = \sum_{j=1}^{m_2} \left\langle \boldsymbol{t}_j^{(k)}, \boldsymbol{\xi} \right\rangle z_j^{(k)} \quad \text{for any} \quad \boldsymbol{\xi} \in \mathcal{H}.
$$

Consequently, we also have

$$
\mathcal{D}_{xx}^2 \rho^{(k)}(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\zeta}) = \sum_{j=1}^{m_2} \left\langle \boldsymbol{t}_j^{(k)}, \boldsymbol{\xi} \right\rangle \left(\mathcal{D}_x \, z_j^{(k)}(\boldsymbol{\zeta}) \right) = \sum_{j=1}^{m_2} \left\langle \boldsymbol{t}_j^{(k)}, \boldsymbol{\xi} \right\rangle \left\langle \nabla_x z_j^{(k)}, \boldsymbol{\zeta} \right\rangle, \quad \text{for any } \boldsymbol{\xi}, \boldsymbol{\zeta} \in \mathcal{H}.
$$

In summary, the first-order and the second-order Fréchet derivatives of $\rho^{(k)}(\cdot,\cdot)$ read

$$
\begin{aligned}
\mathcal{D}_x \rho^{(k)}(\mu, \boldsymbol{x})(\boldsymbol{\xi}) &= \left\langle \mathcal{T}^{(k)\dagger} \boldsymbol{z}^{(k)}, \boldsymbol{\xi} \right\rangle &&= \left(\mathcal{T}^{(k)} \boldsymbol{\xi} \right)^{\mathsf{T}} \boldsymbol{z}^{(k)}, \\
\mathcal{D}_{xx}^2 \rho^{(k)}(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\zeta}) &= \left\langle \mathcal{T}^{(k)\dagger} \left(\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\zeta} \right), \boldsymbol{\xi} \right\rangle &&= \left(\mathcal{T}^{(k)} \boldsymbol{\xi} \right)^{\mathsf{T}} \left(\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\zeta} \right).
\end{aligned} \tag{17}
$$

Therefore, we also need to determine the derivative of $\boldsymbol{z}^{(k)}$ with respect to $\boldsymbol{x}$. Differentiating (15) with respect to $\boldsymbol{x}$, we obtain the system

$$
\begin{aligned}
& S^{(k)} \left(\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\xi} \right) + Z^{(k)} \left(\mathcal{D}_x \boldsymbol{s}^{(k)} \otimes \boldsymbol{\xi} \right) = \mathbf{0}, \\
& \mathcal{W}^{(k)} \mathcal{D}_x \left\langle \boldsymbol{y}^{(k)}, \boldsymbol{\xi} \right\rangle + \mathcal{D}_x \boldsymbol{s}^{(k)} \otimes \boldsymbol{\xi} = -\mathcal{T}^{(k)} \boldsymbol{\xi}, \\
& \left\langle \mathcal{D}_x \left\langle \boldsymbol{y}^{(k)}, \boldsymbol{\xi} \right\rangle, \ \boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger} \boldsymbol{z}^{(k)} \right\rangle + \left\langle \boldsymbol{y}^{(k)}, \ \mathcal{W}^{(k)\dagger} \left(\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\xi} \right) \right\rangle = 0,
\end{aligned} \tag{18}
$$

for any $\boldsymbol{\xi} \in \mathcal{H}$.

Solving the system (18), we get

$$
\begin{aligned}
\mathcal{D}_x \left\langle \boldsymbol{y}^{(k)}, \boldsymbol{\xi} \right\rangle &= -\mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger} Q^{(k)^2} \mathcal{T}^{(k)} \boldsymbol{\xi}, \\
\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\xi} &= Q^{(k)} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}, \\
\mathcal{D}_x \boldsymbol{s}^{(k)} \otimes \boldsymbol{\xi} &= -Q^{(k)^{-1}} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi},
\end{aligned} \tag{19}
$$

for any $\boldsymbol{\xi} \in \mathcal{H}$, where

$$
\begin{aligned}
Q^{(k)} &:= Q^{(k)}(\mu, \boldsymbol{x}) = \left(S^{(k)^{-1}} Z^{(k)} \right)^{\frac{1}{2}}, \\
\mathcal{R}^{(k)} \boldsymbol{\zeta} &:= \mathcal{R}^{(k)}(\mu, \boldsymbol{x}) \boldsymbol{\zeta} = \mathcal{W}^{(k)\dagger} Q^{(k)^2} \mathcal{W}^{(k)} \boldsymbol{\zeta}, \\
P^{(k)} \boldsymbol{v} &:= P^{(k)}(\mu, \boldsymbol{x}) \boldsymbol{v} = \boldsymbol{v} - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger} Q^{(k)} \boldsymbol{v},
\end{aligned} \tag{20}
$$

for any $\boldsymbol{\zeta} \in \mathcal{H}$ and $\boldsymbol{v} \in \mathbb{R}^{m_2}$. Note that, under Assumption 3.1, $\mathcal{R}^{(k)} \cdot$ is an invertible operator from the Hilbert space $\mathcal{H}$ into itself, and hence the operator $\mathcal{R}^{(k)^{-1}} \cdot$ is well-defined on $\mathcal{H}$.

By differentiating $\mathcal{A}\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b}$ with respect to $\boldsymbol{x}$, we get $\mathcal{D}_x \boldsymbol{s} \otimes \boldsymbol{\xi} = -\mathcal{A}\boldsymbol{\xi}$ for any $\boldsymbol{\xi} \in \mathcal{H}$. We then have

$$
\begin{aligned}
\mathcal{D}_x \eta(\mu, \boldsymbol{x}) &= \boldsymbol{c} - \sum_{k=1}^{K} \mathcal{D}_x \rho^{(k)}(\mu, \boldsymbol{x}) + \mu \, \mathcal{D}_x \boldsymbol{s}^{\dagger} \circ \boldsymbol{s}^{-1} \\
&= \boldsymbol{c} - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} \boldsymbol{z}^{(k)} - \mu \, \mathcal{A}^{\dagger} \boldsymbol{s}^{-1},
\end{aligned} \tag{21}
$$

and

$$
\begin{aligned}
\mathcal{D}^2_{xx} \eta(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\zeta}) &= -\sum_{k=1}^{K} \left\langle \mathcal{T}^{(k)\dagger} \left(\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\xi} \right), \boldsymbol{\zeta} \right\rangle + \mu \left\langle \mathcal{A}^{\dagger} S^{-2} \left(\mathcal{D}_x \boldsymbol{s} \otimes \boldsymbol{\xi} \right), \boldsymbol{\zeta} \right\rangle \\
&= -\sum_{k=1}^{K} \left\langle \mathcal{T}^{(k)\dagger} Q^{(k)} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}, \boldsymbol{\zeta} \right\rangle - \mu \left\langle \mathcal{A}^{\dagger} S^{-2} \mathcal{A} \boldsymbol{\xi}, \boldsymbol{\zeta} \right\rangle,
\end{aligned} \tag{22}
$$

for any $\boldsymbol{\xi}, \boldsymbol{\zeta} \in \mathcal{H}$, where $Q^{(k)}$ and $P^{(k)} \cdot$ are defined in (20) for $k = 1, 2, \ldots, K$.

5 Fundamental Properties of the Recourse Function

In this section, we present two fundamental properties of the recourse function $\eta(\cdot, \cdot)$ that yield nice performance of Newton's method used for the proposed algorithms (see [1, Section 1] and [4, Section 3]).

Let $\mathbb{R}_{++}$ denote the set of positive real numbers. For any $\mu > 0$, $\boldsymbol{x} \in \mathcal{F}_0$ and $\boldsymbol{\xi} \in \mathcal{H}$, it is easy to verify the following properties of the function $\eta(\cdot, \cdot)$:

Property 5.1 *Along every sequence $\{x_i \in \mathcal{F}_0\}_{i=1}^{\infty}$ converging to the boundary of $\mathcal{F}_0$, the function $\eta(\mu, x_i)$ tends to infinity.*

Property 5.2 *The function $\eta(\mu, x)$ is continuous on $\mathbb{R}_{++} \times \mathcal{F}_0$, and convex on $\mathcal{F}_0$ for fixed $\mu \in \mathbb{R}_{++}$.*

Property 5.3 *The function $\eta(\mu, x)$ has three Fréchet derivatives on $\mathcal{F}_0$, which are continuous on $\mathbb{R}_{++} \times \mathcal{F}_0$ and continuously differentiable in $\mu \in \mathbb{R}_{++}$.*

Although Properties 5.1, 5.2 and 5.3 of $\eta(\cdot, \cdot)$ are quite important, they are not sufficient to ensure a nice performance of Newton's method used in the proposed algorithms. Some additional features of $\eta(\cdot, \cdot)$ are needed to reach our goal. For this, we first give the following technical result:

Lemma 5.1 *For every $\mu > 0, x \in \mathcal{F}_2^{(k)}$ and $\xi \in \mathcal{H}$, the following inequality holds*

$$\left|\mathcal{D}^3_{xxx}\rho^{(k)}(\mu, x)(\xi, \xi, \xi)\right| \leq 2\mu^{-1/2}\left(\mathcal{D}^2_{xx}\rho^{(k)}(\mu, x)(\xi, \xi)\right)^{3/2}. \tag{23}$$

Proof For any $\mu > 0, x \in \mathcal{F}_2^{(k)}$ and $\xi \in \mathcal{H}$, we define the univariate function

$$\Phi^{(k)}(t) := \mathcal{D}^2_{xx}\rho^{(k)}(\mu, x + t\xi)(\xi, \xi).$$

Note that $\Phi^{(k)}(0) = \mathcal{D}^2_{xx}\rho^{(k)}(\mu, x)(\xi, \xi)$ and $\Phi^{(k)\prime}(0) = \mathcal{D}^3_{xxx}\rho^{(k)}(\mu, x)(\xi, \xi, \xi)$. So, to prove that (23) is satisfied for $\rho^{(k)}(\mu, x)$ on $\mathcal{F}_2^{(k)}$, it suffices to show that

$$\left|\Phi^{(k)\prime}(0)\right| \leq \frac{2}{\sqrt{\mu}}\left|\Phi^{(k)}(0)\right|^{3/2}.$$

Let $\left(y^{(k)}(t), s^{(k)}(t), z^{(k)}(t)\right) := \left(y^{(k)}(\mu, x + t\xi), s^{(k)}(\mu, x + t\xi), z^{(k)}(\mu, x + t\xi)\right)$. For any $\zeta \in \mathcal{H}$ and $v \in \mathbb{R}^{m_2}$, we also let $Q^{(k)}(t) := Q^{(k)}(\mu, x + t\xi), \mathcal{R}^{(k)}(t)\zeta := \mathcal{R}^{(k)}(\mu, x + t\xi)\zeta,$ and $P^{(k)}(t)v := P^{(k)}(\mu, x + t\xi)v$. If we use notations introduced earlier, we have

$$\left(y^{(k)}, s^{(k)}, z^{(k)}\right) = \left(y^{(k)}(0), s^{(k)}(0), z^{(k)}(0)\right), \quad Q^{(k)} = Q^{(k)}(0), \quad \mathcal{R}^{(k)}\zeta = \mathcal{R}^{(k)}(0)\zeta, \quad \text{and} \quad P^{(k)}v = P^{(k)}(0)v.$$

We also define $\boldsymbol{u}^{(k)}(t) := P^{(k)}(t)Q^{(k)}(t)\mathcal{T}^{(k)}(t)\boldsymbol{\xi}$ and $\boldsymbol{u}^{(k)} := \boldsymbol{u}^{(k)}(0)$. It can be easily verified that $P^{(k)^2}\boldsymbol{v} = P^{(k)}\boldsymbol{v}$. Then, using (17) and (19), we have

$$
\begin{aligned}
\Phi^{(k)}(0) &= \mathcal{D}^2_{xx}\rho^{(k)}(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\xi}) \\
&= \left\langle \mathcal{T}^{(k)\dagger}\left(\mathcal{D}_x \boldsymbol{z}^{(k)} \otimes \boldsymbol{\xi}\right), \boldsymbol{\xi}\right\rangle \\
&= \left\langle \mathcal{T}^{(k)\dagger} Q^{(k)} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}, \boldsymbol{\xi}\right\rangle \\
&= \left\langle \mathcal{T}^{(k)\dagger} Q^{(k)} P^{(k)^2} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}, \boldsymbol{\xi}\right\rangle \\
&= \left(P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right)^{\mathsf{T}} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi} \\
&= \left\|\boldsymbol{u}^{(k)}\right\|_2^2,
\end{aligned}
\tag{24}
$$

where the fifth equality follows by applying (2).

Hence $\Phi^{(k)\prime}(0) = 2\boldsymbol{u}^{(k)^{\mathsf{T}}}\boldsymbol{u}^{(k)\prime}$. So, in order to bound $\left|\Phi^{(k)\prime}(0)\right|$, we need to compute the derivative of $\boldsymbol{u}^{(k)}$ with respect to t. Using (20), we have

$$
\begin{aligned}
\boldsymbol{u}^{(k)\prime} &= \left\{P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right\}' \\
&= \left\{Q^{(k)} - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger} Q^{(k)^2}\right\}' \mathcal{T}^{(k)} \boldsymbol{\xi} \\
&= \Big\{Q^{(k)\prime} - Q^{(k)\prime} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger} Q^{(k)^2} + Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger}\left(Q^{(k)} Q^{(k)\prime} + Q^{(k)\prime} Q^{(k)}\right) \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger} Q^{(k)^2} \\
&\qquad - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger}\left(Q^{(k)} Q^{(k)\prime} + Q^{(k)\prime} Q^{(k)}\right)\Big\} \mathcal{T}^{(k)} \boldsymbol{\xi} \\
&= \left\{Q^{(k)\prime} - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger}\left(Q^{(k)} Q^{(k)\prime} + Q^{(k)\prime} Q^{(k)}\right)\right\}\left(I - \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger} Q^{(k)^2}\right) \mathcal{T}^{(k)} \boldsymbol{\xi} \\
&= \left\{Q^{(k)\prime} - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)\dagger}\left(Q^{(k)} Q^{(k)\prime} + Q^{(k)\prime} Q^{(k)}\right)\right\} Q^{(k)^{-1}} \boldsymbol{u}^{(k)}.
\end{aligned}
$$

Since the matrix $Q^{(k)}$ is symmetric, it can also be easily verified that the linear map $P^{(k)}\cdot$ is a self-adjoint operator on $\mathbb{R}^{m_2}$ with the Euclidean inner product. That is, $(P^{(k)}\boldsymbol{v})^{\mathsf{T}}\boldsymbol{w} = \boldsymbol{v}^{\mathsf{T}}P^{(k)}\boldsymbol{w}$ for any $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^{m_2}$. Then, for any $\boldsymbol{\zeta} \in \mathcal{H}$, we have

$$
\begin{aligned}
\boldsymbol{u}^{(k)^{\mathsf{T}}} Q^{(k)} \mathcal{W}^{(k)} \boldsymbol{\zeta} &= \left(P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right)^{\mathsf{T}} Q^{(k)} \mathcal{W}^{(k)} \boldsymbol{\zeta} \\
&= \left(Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right)^{\mathsf{T}} P^{(k)} Q^{(k)} \mathcal{W}^{(k)} \boldsymbol{\zeta} \\
&= \left(Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right)^{\mathsf{T}}\left(Q^{(k)} \mathcal{W}^{(k)} \boldsymbol{\zeta} - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{W}^{(k)^{\mathsf{T}}} Q^{(k)^2} \mathcal{W}^{(k)} \boldsymbol{\zeta}\right) \\
&= \left(Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right)^{\mathsf{T}}\left(Q^{(k)} \mathcal{W}^{(k)} \boldsymbol{\zeta} - Q^{(k)} \mathcal{W}^{(k)} \mathcal{R}^{(k)^{-1}} \mathcal{R}^{(k)} \boldsymbol{\zeta}\right) \\
&= 0.
\end{aligned}
$$

This implies that

$$\Phi^{(k)\prime}(0) = 2\, \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)\prime} = 2\, \boldsymbol{u}^{(k)^{\mathsf{T}}} Q^{(k)\prime} Q^{(k)^{-1}} \boldsymbol{u}^{(k)}. \tag{25}$$

Note that the matrices $Q^{(k)^{-3}}$ and $Q^{(k)\prime}$ are both symmetric (they are, indeed, not only symmetric but also diagonal). Then, by using (19), (20), (24), (25) and norm inequalities, we get

$$\begin{aligned}
\left|\Phi^{(k)\prime}(0)\right| &= 2\, \left|\boldsymbol{u}^{(k)^{\mathsf{T}}} Q^{(k)\prime} Q^{(k)^{-1}} \boldsymbol{u}^{(k)}\right| \\
&= \left|\boldsymbol{u}^{(k)^{\mathsf{T}}} \left(Q^{(k)\prime} Q^{(k)^{-1}} + Q^{(k)^{-1}} Q^{(k)\prime}\right) \boldsymbol{u}^{(k)}\right| \\
&= \left|\boldsymbol{u}^{(k)^{\mathsf{T}}} Q^{(k)} \left(Q^{(k)^{-3}} Q^{(k)\prime} + Q^{(k)\prime} Q^{(k)^{-3}}\right) Q^{(k)} \boldsymbol{u}^{(k)}\right| \\
&= \left|\boldsymbol{u}^{(k)^{\mathsf{T}}} Q^{(k)} \left(Q^{(k)^{-2}}\right)' Q^{(k)} \boldsymbol{u}^{(k)}\right| \\
&\leq \left\|Z^{(k)^2} \left(Q^{(k)^{-2}}\right)' \boldsymbol{e}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} Q^{(k)} Z^{(k)^{-2}} Q^{(k)} \boldsymbol{u}^{(k)} \\
&\leq \mu^{-1} \left\|Z^{(k)^2} \left(Q^{(k)^{-2}}\right)' \boldsymbol{e}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&= \mu^{-1} \left\|Z^{(k)^2} \left(Z^{(k)^{-1}} S^{(k)}\right)' \boldsymbol{e}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&= \mu^{-1} \left\|Z^{(k)} \boldsymbol{s}^{(k)\prime} - S^{(k)} \boldsymbol{z}^{(k)\prime}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&= \mu^{-1} \left\|Z^{(k)} \left(\mathcal{D}_x \boldsymbol{s}^{(k)}(\mu, \boldsymbol{x} + t\boldsymbol{\xi}) \otimes \boldsymbol{\xi}\right) - S^{(k)} \left(\mathcal{D}_x \boldsymbol{z}^{(k)}(\mu, \boldsymbol{x} + t\boldsymbol{\xi}) \otimes \boldsymbol{\xi}\right)\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&= \mu^{-1} \left\|Z^{(k)} Q^{(k)^{-1}} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi} + S^{(k)} Q^{(k)} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&= \mu^{-1} \left\|\left(Z^{(k)} Q^{(k)^{-1}} + S^{(k)} Q^{(k)}\right) P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&= 2\mu^{-1/2} \left\|P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right\|_\infty \boldsymbol{u}^{(k)^{\mathsf{T}}} \boldsymbol{u}^{(k)} \\
&\leq 2\mu^{-1/2} \left\|P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi}\right\|_2 \left\|\boldsymbol{u}^{(k)}\right\|_2^2 \\
&= 2\mu^{-1/2} \left|\Phi^{(k)}(0)\right|^{3/2}.
\end{aligned}$$

The proof is complete. □

The following theorem gives a significant result. This result will deduce the first fundamental property of the recourse function. It will be also employed in the proofs of the iteration complexity theorems of the proposed algorithms.

Theorem 5.1 *For every $\mu > 0, \boldsymbol{x} \in \mathcal{F}_0$ and $\boldsymbol{\xi} \in \mathcal{H}$, the following inequality holds*

$$\left|\mathcal{D}^3_{xxx} \eta(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\xi}, \boldsymbol{\xi})\right| \leq 2\mu^{-1/2} \left(\mathcal{D}^2_{xx} \eta(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\xi})\right)^{3/2}. \tag{26}$$

Proof Let $\mu > 0, \boldsymbol{x} \in \mathcal{F}_0$ and $\boldsymbol{\xi} \in \mathcal{H}$. Let also $\ell(\mu, \boldsymbol{x}) := -\mu\, \boldsymbol{e}^{\mathsf{T}} \ln \boldsymbol{s}(\mu, \boldsymbol{x})$. Then

$$\eta(\mu, \boldsymbol{x}) = \langle \boldsymbol{c}, \boldsymbol{x} \rangle + \ell(\mu, \boldsymbol{x}) + \sum_{k=1}^{K} \rho^{(k)}(\mu, \boldsymbol{x}).$$

Note that the inequality in (26) is not affected by the map $\langle \boldsymbol{c}, \boldsymbol{x} \rangle$ due to its linearity (the Hessian and the third-order derivative of the linear map $\langle \boldsymbol{c}, \boldsymbol{x} \rangle$ are identically zeros).

On the other hand, it is easy to verify that the function $\ell(\cdot, \cdot)$ satisfies the following inequality on $\mathcal{F}_1$

$$\left| \mathcal{D}^3_{xxx} \ell(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\xi}, \boldsymbol{\xi}) \right| \leq 2\mu^{-1/2} \left(\mathcal{D}^2_{xx} \ell(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\xi}) \right)^{3/2}.$$

In Lemma 5.1, we have also shown that the function $\rho^{(k)}(\cdot, \cdot)$ satisfies the inequality in (23) on $\mathcal{F}_2^{(k)}$. Following Proposition 2.1.1(ii) in [9], we conclude that the recourse function $\eta(\cdot, \cdot)$ satisfies the inequality in (26) on $\mathcal{F}_0$. The proof is complete. □

The fundamental property presented below plays a central role in developing interior-point methods for our problem setting [1, 9].

Fundamental Property 5.1 *Property* 5.1 *and Theorem* 5.1 *mean that, for any fixed* $\mu > 0$, *the recourse function* $\eta(\cdot, \cdot)$ *is a* μ *strongly self-concordant function on* $\mathcal{F}_0$. *(For the complete and precise definition of strongly self-concordant functionals, see for example, [9, Definition 2.1.1] for the finite-dimensional case and [1, Section 1] for the infinite-dimensional case).*

To arrive at the second fundamental property of $\eta(\cdot, \cdot)$, we first prove the following two lemmas.

Lemma 5.2 *For any* $\mu > 0, \boldsymbol{x} \in \mathcal{F}_0$ *and* $\boldsymbol{\xi} \in \mathcal{H}$, *the following inequality holds*

$$\left| \{ \mathcal{D}_x \eta(\mu, \boldsymbol{x})(\boldsymbol{\xi}) \}' \right| \leq \left(-\frac{m_1 + K m_2}{\mu}\, \mathcal{D}^2_{xx} \eta(\mu, \boldsymbol{x})(\boldsymbol{\xi}, \boldsymbol{\xi}) \right)^{1/2}.$$

Proof By differentiating (15) with respect to μ, we obtain the system

$$\begin{aligned} S^{(k)} \boldsymbol{z}^{(k)\prime} + Z^{(k)} \boldsymbol{s}^{(k)\prime} &= \boldsymbol{e}, \\ \mathcal{W}^{(k)} \boldsymbol{y}^{(k)\prime} + \boldsymbol{s}^{(k)\prime} &= \boldsymbol{0}, \\ \left\langle \boldsymbol{y}^{(k)\prime},\ \boldsymbol{d}^{(k)} + \mathcal{W}^{(k)\dagger} \boldsymbol{z}^{(k)} \right\rangle + \left\langle \boldsymbol{y}^{(k)},\ \mathcal{W}^{(k)\dagger} \boldsymbol{z}^{(k)\prime} \right\rangle &= 0. \end{aligned} \tag{27}$$

After solving (27), we get

$$\begin{aligned} \boldsymbol{y}^{(k)'} &= -\mathcal{R}^{(k)^{-1}}\mathcal{W}^{(k)\dagger}\boldsymbol{s}^{(k)^{-1}}, \\ \boldsymbol{z}^{(k)'} &= \frac{1}{\sqrt{\mu}} Q^{(k)} P^{(k)} \boldsymbol{e}, \\ \boldsymbol{s}^{(k)'} &= \mathcal{W}^{(k)}\mathcal{R}^{(k)^{-1}}\mathcal{W}^{(k)\dagger}\boldsymbol{s}^{(k)^{-1}}. \end{aligned} \tag{28}$$

Differentiating (21) with respect to μ and applying (28), we obtain

$$\{\mathcal{D}_x \eta(\mu, \boldsymbol{x})\}' = -\frac{1}{\sqrt{\mu}} \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} Q^{(k)} P^{(k)} \boldsymbol{e} - \mathcal{A}^{\dagger} S^{-1} \boldsymbol{e}.$$

We now define

$$\begin{aligned} \mathcal{M}\boldsymbol{\xi} &:= -S^{-1}\mathcal{A}\boldsymbol{\xi} \in \mathbb{R}^{m_1}, & \mathcal{M}^{\dagger}\boldsymbol{\vartheta} &:= -\mathcal{A}^{\dagger}S^{-1}\boldsymbol{\vartheta} \in \mathcal{H}, \\ \mathcal{M}^{(k)}\boldsymbol{\xi} &:= -\frac{1}{\sqrt{\mu}} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi} \in \mathbb{R}^{m_2}, & \mathcal{M}^{(k)\dagger}\boldsymbol{v} &:= -\frac{1}{\sqrt{\mu}} \mathcal{T}^{(k)\dagger} Q^{(k)} P^{(k)} \boldsymbol{v} \in \mathcal{H}, \end{aligned}$$

and

$$\mathfrak{M}\boldsymbol{\xi} := \sum_{k=1}^{K} \mathcal{M}^{(k)\dagger}\mathcal{M}^{(k)}\boldsymbol{\xi} + \mathcal{M}^{\dagger}\mathcal{M}\boldsymbol{\xi} = \frac{1}{\mu}\sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} Q^{(k)} P^{(k)} Q^{(k)} \mathcal{T}^{(k)} \boldsymbol{\xi} + \mathcal{A}^{\dagger} S^{-2} \mathcal{A}\boldsymbol{\xi},$$

for any $\boldsymbol{\xi} \in \mathcal{H}$, $\boldsymbol{\vartheta} \in \mathbb{R}^{m_1}$ and $\boldsymbol{v} \in \mathbb{R}^{m_2}$. Then, using (2), we have

$$\begin{aligned} \langle \mathfrak{M}\boldsymbol{\xi}, \boldsymbol{\xi} \rangle &= \sum_{k=1}^{K} \left\langle \mathcal{M}^{(k)\dagger}\mathcal{M}^{(k)}\boldsymbol{\xi}, \boldsymbol{\xi} \right\rangle - \mu \left\langle \mathcal{M}^{\dagger}\mathcal{M}\boldsymbol{\xi}, \boldsymbol{\xi} \right\rangle \\ &= \sum_{k=1}^{K} \left(\mathcal{M}^{(k)}\boldsymbol{\xi}\right)^{\mathsf{T}} \mathcal{M}^{(k)}\boldsymbol{\xi} - \mu \left(\mathcal{M}\boldsymbol{\xi}\right)^{\mathsf{T}} \mathcal{M}\boldsymbol{\xi} \\ &= (\Psi\boldsymbol{\xi})^{\mathsf{T}} \Psi\boldsymbol{\xi} \\ &= \left\langle \Psi^{\dagger}\Psi\boldsymbol{\xi}, \boldsymbol{\xi} \right\rangle, \end{aligned}$$

where, for $\boldsymbol{\xi} \in \mathcal{H}, \boldsymbol{\vartheta} \in \mathbb{R}^{m_1}$ and $\boldsymbol{v}_k \in \mathbb{R}^{m_2}, k = 1, 2, \ldots, K$, $\Psi\boldsymbol{\cdot}$ and $\Psi^{\dagger}\boldsymbol{\cdot}$ are defined as

$$\Psi\boldsymbol{\xi} := \begin{bmatrix} \mathcal{M}^{(1)}\boldsymbol{\xi} \\ \vdots \\ \mathcal{M}^{(K)}\boldsymbol{\xi} \\ \mathcal{M}\boldsymbol{\xi} \end{bmatrix} \in \underbrace{\mathbb{R}^{m_2} \times \cdots \times \mathbb{R}^{m_2}}_{K-\text{times}} \times \mathbb{R}^{m_1}, \text{ and } \Psi^{\dagger} \begin{bmatrix} \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{v}_K \\ \boldsymbol{\vartheta} \end{bmatrix} := \sum_{k=1}^{K} \mathcal{M}^{(k)\dagger}\boldsymbol{v}_k + \mathcal{M}^{\dagger}\boldsymbol{\vartheta} \in \mathcal{H}.$$

Note that the operator $\mathcal{M}\cdot = \Psi^{\dagger}\Psi\cdot$ is invertible from the Hilbert space $\mathcal{H}$ into itself, and hence its inverse operator $\mathcal{M}^{-1}\cdot = \left(\Psi^{\dagger}\Psi\right)^{-1}\cdot$ is well-defined on $\mathcal{H}$.

Note also that

$$\mathcal{D}^2_{xx}\eta(\mu,x)(\xi,\xi) = -\mu\,\langle \mathcal{M}\xi,\xi\rangle = -\mu\,(\Psi\xi)^{\mathsf{T}}\,\Psi\xi = -\mu\left\langle \Psi^{\dagger}\Psi\xi,\xi\right\rangle, \tag{29}$$

and that

$$\{\mathcal{D}_x\eta(\mu,x)(\xi)\}' = \sum_{k=1}^{K}\left\langle \mathcal{M}^{(k)\dagger}e,\xi\right\rangle + \left\langle \mathcal{M}^{\dagger}e,\xi\right\rangle = \sum_{k=1}^{K} e^{\mathsf{T}}\mathcal{M}^{(k)}\xi + e^{\mathsf{T}}\mathcal{M}\xi = \varepsilon^{\mathsf{T}}\Psi\xi = \left\langle \Psi^{\dagger}\varepsilon,\xi\right\rangle, \tag{30}$$

where

$$\varepsilon := \begin{bmatrix} e \\ \vdots \\ e \\ e \end{bmatrix} \in \underbrace{\mathbb{R}^{m_2}\times\cdots\times\mathbb{R}^{m_2}}_{K-\text{times}}\times\mathbb{R}^{m_1}.$$

By (29) and (30), we get

$$\begin{aligned} -\left\{\mathcal{D}^2_{xx}\eta(\mu,x)\right\}^{-1}\left(\{\mathcal{D}_x\eta(\mu,x)\}',\{\mathcal{D}_x\eta(\mu,x)\}'\right) &= \frac{1}{\mu}\left\langle\left(\Psi^{\dagger}\Psi\right)^{-1}\Psi^{\dagger}\varepsilon,\{\nabla_x\eta(\mu,x)\}'\right\rangle \\ &= \frac{1}{\mu}\varepsilon^{\mathsf{T}}\Psi\left(\Psi^{\dagger}\Psi\right)^{-1}\Psi^{\dagger}\varepsilon \\ &\le \frac{1}{\mu}\varepsilon^{\mathsf{T}}\varepsilon \\ &= \frac{1}{\mu}\left(m_1 + Km_2\right). \end{aligned} \tag{31}$$

Using norm inequalities and (31), we have

$$\begin{aligned} \left|\{\mathcal{D}_x\eta(\mu,x)(\xi)\}'\right| &\le \sqrt{-\left\{\mathcal{D}^2_{xx}\eta(\mu,x)\right\}^{-1}\left(\{\mathcal{D}_x\eta(\mu,x)\}',\{\mathcal{D}_x\eta(\mu,x)\}'\right)}\sqrt{-\mathcal{D}^2_{xx}\eta(\mu,x)(\xi,\xi)} \\ &\le \sqrt{-\frac{m_1+Km_2}{\mu}\,\mathcal{D}^2_{xx}\eta(\mu,x)(\xi,\xi)}. \end{aligned}$$

The proof is complete. □

Lemma 5.3 *For any $\mu > 0, x \in \mathcal{F}_0$ and $\xi \in \mathcal{H}$, the following inequality holds*

$$\left|\left\{\mathcal{D}^2_{xx}\eta(\mu,x)(\xi,\xi)\right\}'\right| \le -\frac{\sqrt{m_2}}{\mu}\,\mathcal{D}^2_{xx}\eta(\mu,x)(\xi,\xi).$$

Proof Let $\left(\boldsymbol{s}^{(k)}, \boldsymbol{z}^{(k)}, Q^{(k)}, \mathcal{R}^{(k)}, P^{(k)}\right) := \left(\boldsymbol{s}^{(k)}(\mu,\boldsymbol{x}), \boldsymbol{z}^{(k)}(\mu,\boldsymbol{x}), Q^{(k)}(\mu,\boldsymbol{x}), \mathcal{R}^{(k)}(\mu,\boldsymbol{x}), P^{(k)}(\mu,\boldsymbol{x})\right)$. We fix $\boldsymbol{\xi} \in \mathcal{H}$ and define $\boldsymbol{u}^{(k)} := P^{(k)}Q^{(k)}\mathcal{T}^{(k)}\boldsymbol{\xi}$. Then we have

$$\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})(\boldsymbol{\xi},\boldsymbol{\xi}) = -\sum_{k=1}^{K} \boldsymbol{u}^{(k)^\mathsf{T}}\boldsymbol{u}^{(k)} - \mu\left\langle \mathcal{A}^\dagger S^{-2}\mathcal{A}\boldsymbol{\xi},\boldsymbol{\xi}\right\rangle.$$

Following the steps in Lemma 5.1 leading up to the fourth equality in (5), we get

$$\left\{\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})(\boldsymbol{\xi},\boldsymbol{\xi})\right\}' = -\sum_{k=1}^{K} \boldsymbol{u}^{(k)^\mathsf{T}} Q^{(k)}\left(Q^{(k)^{-2}}\right)' Q^{(k)}\boldsymbol{u}^{(k)} - \left\langle \mathcal{A}^\dagger S^{-2}\mathcal{A}\boldsymbol{\xi},\boldsymbol{\xi}\right\rangle.$$

Using (20), we have

$$\begin{aligned}
\boldsymbol{u}^{(k)^\mathsf{T}} Q^{(k)}\left(Q^{(k)^{-2}}\right)' Q^{(k)}\boldsymbol{u}^{(k)} &= \boldsymbol{u}^{(k)^\mathsf{T}}\left(S^{(k)^{-1}}Z^{(k)}\right)^{\frac{1}{2}}\left(Z^{(k)^{-1}}S^{(k)}\right)'\left(S^{(k)^{-1}}Z^{(k)}\right)^{\frac{1}{2}}\boldsymbol{u}^{(k)}\\
&= \boldsymbol{u}^{(k)^\mathsf{T}}\left(\mu^{-1}S^{(k)^2}\right)^{-\frac{1}{2}}\left(Z^{(k)^{-1}}S^{(k)}\right)'\left(\mu^{-1}S^{(k)^2}\right)^{-\frac{1}{2}}\boldsymbol{u}^{(k)}\\
&= \mu\,\boldsymbol{u}^{(k)^\mathsf{T}} S^{(k)^{-1}}\left(Z^{(k)^{-1}}S^{(k)'} - S^{(k)}Z^{(k)^{-2}}Z^{(k)'}\right)S^{(k)^{-1}}\boldsymbol{u}^{(k)}\\
&= \mu^{-1}\,\boldsymbol{u}^{(k)^\mathsf{T}}\left(Z^{(k)}S^{(k)'} - S^{(k)}Z^{(k)'}\right)\boldsymbol{u}^{(k)}\\
&= \mu^{-1}\,\boldsymbol{u}^{(k)^\mathsf{T}}\left(2\,Z^{(k)}S^{(k)'} - I\right)\boldsymbol{u}^{(k)}\\
&= \mu^{-1}\,\boldsymbol{u}^{(k)^\mathsf{T}}\left(2\mu\,S^{(k)^{-1}}S^{(k)'} - I\right)\boldsymbol{u}^{(k)}\\
&\leq \mu^{-1}\left\|\boldsymbol{u}^{(k)}\right\|_2^2\left\|\boldsymbol{e} - 2\mu\,S^{(k)^{-1}}\boldsymbol{s}^{(k)'}\right\|_2\\
&= \mu^{-1}\left\|\boldsymbol{u}^{(k)}\right\|_2^2\left\|\boldsymbol{e} - 2\mu\,S^{(k)^{-1}}\mathcal{W}^{(k)}\mathcal{R}^{(k)^{-1}}\mathcal{W}^{(k)\dagger}\boldsymbol{s}^{(k)^{-1}}\right\|_2\\
&= \mu^{-1}\left\|\boldsymbol{u}^{(k)}\right\|_2^2\left\|\left(I - 2\,P^{(k)}\right)\boldsymbol{e}\right\|_2\\
&\leq \frac{\sqrt{m_2}}{\mu}\left\|\boldsymbol{u}^{(k)}\right\|_2^2,
\end{aligned}$$

where the fifth equality follows from the first equation in (27), the seventh equality follows from the third equation in (28), and the last inequality follows from the fact that $I - 2P^{(k)} \leq I$ (hence $\|I - 2\,P^{(k)}\|_2 \leq 1$). It immediately follows that

$$\begin{aligned}
\left|\left\{\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})(\boldsymbol{\xi},\boldsymbol{\xi})\right\}'\right| &\leq \frac{\sqrt{m_2}}{\mu}\sum_{k=1}^{K}\boldsymbol{u}^{(k)^\mathsf{T}}\boldsymbol{u}^{(k)} + \left\langle \mathcal{A}^\dagger S^{-2}\mathcal{A}\boldsymbol{\xi},\boldsymbol{\xi}\right\rangle\\
&\leq \frac{\sqrt{m_2}}{\mu}\left(\sum_{k=1}^{K}\boldsymbol{u}^{(k)^\mathsf{T}}\boldsymbol{u}^{(k)} + \mu\,\left\langle \mathcal{A}^\dagger S^{-2}\mathcal{A}\boldsymbol{\xi},\boldsymbol{\xi}\right\rangle\right)\\
&= -\frac{\sqrt{m_2}}{\mu}\,\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})(\boldsymbol{\xi},\boldsymbol{\xi}).
\end{aligned}$$

□

We now present the second fundamental property of the recourse function.

Fundamental Property 5.2 *Properties* 5.2 *and* 5.3*, Fundamental Property* 5.1*, and Lemmas* 5.2 *and* 5.3 *mean that the family* $\{\eta(\mu,\cdot) : \mu > 0\}$ *is a strongly self-concordant family with the following parameters*

$$\alpha_1(\mu) = \mu,\ \alpha_2(\mu) = \alpha_3(\mu) = 1,\ \alpha_4(\mu) = \frac{\sqrt{m_1 + Km_2}}{\mu},\ \alpha_5(\mu) = \frac{\sqrt{m_2}}{2\mu}.$$

(For the complete and precise definition of a strongly self-concordant family, see [9, Definition 3.1.1]).

Fundamental Property 5.2 means that we can generalize logarithmic-barrier decomposition interior-point techniques given by Zhao [10], Alzalg [13], Mehrotra and Özevin [15], Alzalg and Ariyawansa [17], and Zhao [18] for the finite-dimensional setting to the infinite-dimensional setting. We will see that the obtained complexity estimates are similar to the finite-dimensional ones, which is not surprising due to the remarkable agreement between the nature of the parameters obtained in Fundamental Property 5.2 and that obtained in [10, 13, 15, 17] for finite-dimensional cases.

6 Logarithmic-Barrier Decomposition Algorithms and Complexity

In Section 4, we computed the Fréchet derivatives $\mathcal{D}_x\eta(\mu,\boldsymbol{x})$ and $\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})$ for the *Newton direction* given by

$$\Delta\boldsymbol{x} := -\left\{\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})\right\}^{-1}\left(\mathcal{D}_x\eta(\mu,\boldsymbol{x}) - \boldsymbol{\tau}^\star(\boldsymbol{x})\right) \tag{32}$$

for our algorithms, where $\boldsymbol{\tau}^\star(\boldsymbol{x})$ is a unique vector in $\mathcal{G}^\perp$ such that $\Delta\boldsymbol{x} \in \mathcal{G}$.

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

$$\delta(\mu,\boldsymbol{x}) := \sqrt{-\frac{1}{\mu}\ \mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})(\Delta\boldsymbol{x},\Delta\boldsymbol{x})}. \tag{33}$$

Note that $\boldsymbol{\tau}^\star(\cdot)$ in (32) is the optimal solution to the minimization problem

$$\begin{aligned} &\min\ \left\{\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})\right\}^{-1}\left(\mathcal{D}_x\eta(\mu,\boldsymbol{x}) - \boldsymbol{\tau}, \mathcal{D}_x\eta(\mu,\boldsymbol{x}) - \boldsymbol{\tau}\right) \\ &\text{s.t.}\quad \boldsymbol{\tau} \in \mathcal{G}^\perp. \end{aligned} \tag{34}$$

By Theorem 2.1, the minimization problem (34) has a unique optimal solution, which is characterized by

$$\left\{\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})\right\}^{-1}(\mathcal{D}_x\eta(\mu,\boldsymbol{x})-\bar{\boldsymbol{\tau}})\in\mathcal{G},$$

and hence $\boldsymbol{\tau}^\star(\boldsymbol{x})=\bar{\boldsymbol{\tau}}\in\mathcal{G}^\perp$.

Note also that $\delta(\cdot,\cdot)$ in (33) vanishes at $(\mu,\boldsymbol{x})$, if and only if,

$$\left(\boldsymbol{x},\boldsymbol{s};\boldsymbol{y}^{(1)},\boldsymbol{s}_2^{(1)};\ldots;\boldsymbol{y}^{(K)},\boldsymbol{s}^{(K)}\right)=\left(\boldsymbol{x}(\mu),\boldsymbol{s}(\mu);\boldsymbol{y}^{(1)}(\mu),\boldsymbol{s}_2^{(1)}(\mu);\ldots;\boldsymbol{y}^{(K)}(\mu),\boldsymbol{s}^{(K)}(\mu)\right),$$

provided that $\left(\boldsymbol{x},\boldsymbol{s};\boldsymbol{y}^{(1)},\boldsymbol{s}_2^{(1)};\ldots;\boldsymbol{y}^{(K)},\boldsymbol{s}^{(K)}\right)$ is feasible for (16).

In this section, we present logarithmic-barrier path-following interior-point decomposition algorithms for solving two-stage stochastic linear programming in a Hilbert space and present the complexity results. This class of algorithms is formally stated in Algorithm 6.1.

Algorithm 6.1: Path-Following Interior-Point Decomposition Algorithms for Problem (12, 13)

Begin algorithm

Initialize $\epsilon,\gamma,\theta,\beta,\boldsymbol{x}^0,\mu^0$

Ensure: $\epsilon>0$, $\gamma\in(0,1)$, $\theta>0$, $\beta>0$, $\boldsymbol{x}^0\in\mathcal{F}_0$, $\mu^0>0$

Set $\boldsymbol{x}:=\boldsymbol{x}^0$, $\mu:=\mu^0$

While $\mu\geq\epsilon$ **do**

For $k=1,2,\ldots,K$ **do**

Solve (15) to obtain $\left(\boldsymbol{y}^{(k)},\boldsymbol{s}^{(k)},\boldsymbol{z}^{(k)}\right)$

End for

Compute $\mathcal{D}_x\eta(\mu,\boldsymbol{x})$ and $\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})$ using (21) and (22)

Compute $\boldsymbol{\tau}^\star(\boldsymbol{x})$ by solving (34)

Compute $\Delta\boldsymbol{x}:=-\left\{\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})\right\}^{-1}(\mathcal{D}_x\eta(\mu,\boldsymbol{x})-\boldsymbol{\tau}^\star(\boldsymbol{x}))$

Compute $\delta(\mu,\boldsymbol{x}):=\sqrt{-\frac{1}{\mu}\mathcal{D}^2_{xx}\eta(\mu,\boldsymbol{x})\,(\Delta\boldsymbol{x},\Delta\boldsymbol{x})}$

While $\delta > \beta$ **do**

Set $\boldsymbol{x} := \boldsymbol{x} + \theta \Delta \boldsymbol{x}$

For $k = 1, 2, \ldots, K$ **do**

Solve (15) to obtain $\left(\boldsymbol{y}^{(k)}, \boldsymbol{s}^{(k)}, \boldsymbol{z}^{(k)}\right)$

End for

Compute $\mathcal{D}_{x}\eta(\mu, \boldsymbol{x})$ and $\mathcal{D}^2_{xx}\eta(\mu, \boldsymbol{x})$ using (21) and (22)

Compute $\boldsymbol{\tau}^{\star}(\boldsymbol{x})$ by solving (34)

Compute $\Delta \boldsymbol{x} := -\left\{\mathcal{D}^2_{xx}\eta(\mu, \boldsymbol{x})\right\}^{-1} \left(\mathcal{D}_{x}\eta(\mu, \boldsymbol{x}) - \boldsymbol{\tau}^{\star}(\boldsymbol{x})\right)$

Compute $\delta(\mu, \boldsymbol{x}) := \sqrt{-\frac{1}{\mu}\mathcal{D}^2_{xx}\eta(\mu, \boldsymbol{x})\,(\Delta \boldsymbol{x}, \Delta \boldsymbol{x})}$

End while

Set $\mu := \gamma\mu$

End while

End algorithm

We initialize Algorithm 6.1 with a starting point $\boldsymbol{x}^0 \in \mathcal{F}_0$ and a starting value $\mu^0 > 0$ for the barrier parameter μ, and index it by a parameter $\gamma \in (0, 1)$. We use δ as a measure of the proximity of the current point $\boldsymbol{x}$ to the central path, and β as a threshold for that measure. If the current $\boldsymbol{x}$ is too far away from the central path in the sense that $\delta > \beta$, we apply Newton's method to find a point close to the central path. Then the value of μ is reduced by a factor γ and the whole process is repeated until the value of μ is within the tolerance ϵ. When we trace the central path as μ approaches zero, we can generate a strictly feasible ϵ-optimal solution to Problem (12).

We have two variants of algorithms, namely, the short-step algorithm and the long-step algorithm, which are based on the selection of γ in Algorithm 6.1. In the short-step version of the algorithm, the barrier parameter is decreased by a factor $\gamma := 1 - \sigma/\sqrt{r_1 + Kr_2}$, with $\sigma < 0.1$, in each iteration. The k^{th} iteration of the short-step algorithm is performed as follows: At the beginning of the iteration, we have $\mu^{(k-1)}$ and $\boldsymbol{x}^{(k-1)}$ on hand and $\boldsymbol{x}^{(k-1)}$ is close to the central path, i.e., $\delta(\mu^{(k-1)}, \boldsymbol{x}^{(k-1)}) \leq \beta$. After the barrier parameter μ is reduced from $\mu^{(k-1)}$ to $\mu^k := \gamma\mu^{(k-1)}$, we have that $\delta(\mu^k, \boldsymbol{x}^{(k-1)}) \leq 2\beta$. Then a full Newton step with size $\theta = 1$

is taken to produce a new point $\boldsymbol{x}^k$ with $\delta(\mu^k, \boldsymbol{x}^k) \leq \beta$. We now show that, in this class of algorithms, only one Newton step is sufficient for recentering after updating the parameter μ.

In the long-step version of the 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^{th} iteration of the long-step algorithms is performed as follows: At the beginning of the iteration we have a point $\boldsymbol{x}^{(k-1)}$, which is sufficiently close to $\boldsymbol{x}(\mu^{(k-1)})$, where $\boldsymbol{x}(\mu^{(k-1)})$ is the solution to (12) for $\mu := \mu^{(k-1)}$. The barrier parameter is reduced from $\mu^{(k-1)}$ to $\mu^k := \gamma\mu^{(k-1)}$, where $\gamma \in (0, 1)$, and then a search is started to find a point $\boldsymbol{x}^k$ that is sufficiently close to $\boldsymbol{x}(\mu^k)$. The long-step algorithm generates a finite sequence consisting of N points in $\mathcal{F}_0$, and we finally take $\boldsymbol{x}^k$ to be equal to the last point of this sequence. We want to determine an upper bound on N, which is the number of Newton iterations needed to find the point $\boldsymbol{x}^k$.

Theorems 6.1 and 6.2 below present the complexity results for the short- and long-step algorithms. The proofs of Theorems 6.1 and 6.2 are similar to the proofs of Theorems 5.3 and 5.7, respectively, in [10] for stochastic linear programming in the finite-dimensional setting, and use Theorem 5.1 which we already established for the infinite-dimensional setting. We only indicate that these proofs make use of Propositions 3.2 and 3.4 and Theorem 3.4 in [4] for the infinite-dimensional setting instead of Theorems 2.2.3 and 3.1.1 in [9] for the finite-dimensional one.

Theorem 6.1 *Let μ^0 be the initial barrier parameter, $\epsilon > 0$ the stopping criterion, and $\beta = (2 - \sqrt{3})/2$. If the starting point $\boldsymbol{x}^0$ is sufficiently close to the central path, i.e., $\delta(\mu^0, \boldsymbol{x}^0) \leq \beta$, then the short-step algorithm reduces the barrier parameter μ at a linear rate and terminates with at most $O\left(\sqrt{m_1 + Km_2}\ \ln(\mu^0/\epsilon)\right)$ iterations.*

Theorem 6.2 *Let μ^0 be the initial barrier parameter, $\epsilon > 0$ the stopping criterion, and $\beta = 1/6$. If the starting point $\boldsymbol{x}^0$ is sufficiently close to the central path, i.e., $\delta(\mu^0, \boldsymbol{x}^0) \leq \beta$, then the long-step algorithm reduces the barrier parameter μ at a linear rate and terminates with at most $O\left((m_1 + Km_2)\ \ln(\mu^0/\epsilon)\right)$ iterations.*

Note that the complexity results obtained in Theorems 6.1 and 6.2 coincide with the complexity results obtained in [10, Theorems 5.3 and 5.7] (see Table 1) for stochastic linear programming which are the best known in the finite-dimensional case. Most notably, our complexity results are independent on the choice of the underlying Hilbert space. From this stems the importance of the development in this paper. Note also

Table 1: Complexity results for different classes of stochastic programming problems.

Two-stage stochastic optimization problem	The first- and second-stage feasibility sets $\mathcal{F}_1$ and $\mathcal{F}_2^{(k)}(\boldsymbol{x})$, respectively	Complexity for short- & long-step algorithms
Stochastic linear programming [10]	$\mathcal{F}_1 = \{\boldsymbol{x} \in \mathbb{R}_+^{m_1} : A\boldsymbol{x} = \boldsymbol{b}\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in \mathbb{R}_+^{m_2} : W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{h}^{(k)} - T^{(k)}\boldsymbol{x}\}$	$O\left(\sqrt{m_1 + Km_1}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((m_1 + Km_2)\ \ln(\frac{\mu^0}{\epsilon})\right)$
Stochastic convex quadratic programming [11]	$\mathcal{F}_1 = \{\boldsymbol{x} \in \mathbb{R}_+^{m_1} : A\boldsymbol{x} = \boldsymbol{b}\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in \mathbb{R}_+^{m_2} : W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{h}^{(k)} - T^{(k)}\boldsymbol{x}\}$	$O\left(\sqrt{m_1 + Km_2}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((m_1 + Km_2)\ \ln(\frac{\mu^0}{\epsilon})\right)$
Stochastic second-order cone programming [13]	$\mathcal{F}_1 = \{\boldsymbol{x} \in \boldsymbol{Q}^{n_1} \times \cdots \times \boldsymbol{Q}^{n_p} \subset \mathbb{R}^{\sum_{i=1}^{p} n_i} : A\boldsymbol{x} = \boldsymbol{b}\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in \boldsymbol{Q}^{m_1} \times \cdots \times \boldsymbol{Q}^{m_q} \subset \mathbb{R}^{\sum_{j=1}^{q} m_j} : W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{h}^{(k)} - T^{(k)}\boldsymbol{x}\}$	$O\left(\sqrt{p + Kq}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((p + Kq)\ \ln(\frac{\mu^0}{\epsilon})\right)$
Stochastic semidefinite programming [15]	$\mathcal{F}_1 = \{\boldsymbol{x} \in \mathbb{R}^{m_1} : B - \sum_{i=1}^{m_1} x_i A_i \in \mathcal{S}^p\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in \mathbb{R}^{m_2} : H^{(k)} - \sum_{i=1}^{m_1} x_i T_i^{(k)} - \sum_{j=1}^{m_2} y_j^{(k)} W_j^{(k)} \in \mathcal{S}^q\}$	$O\left(\sqrt{p + Kq}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((p + Kq)\ \ln(\frac{\mu^0}{\epsilon})\right)$
Stochastic symmetric cone programming [17]	$\mathcal{F}_1 = \{\boldsymbol{x} \in \mathbb{R}^{m_1} : \boldsymbol{b} - A\boldsymbol{x} \in \mathcal{K}^{r_1}\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in \mathbb{R}^{m_2} : \boldsymbol{h}^{(k)} - T^{(k)}\boldsymbol{x} - W^{(k)}\boldsymbol{y}^{(k)} \in \mathcal{K}^{r_2}\}$	$O\left(\sqrt{r_1 + Kr_2}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((r_1 + Kr_2)\ \ln(\frac{\mu^0}{\epsilon})\right)$
Stochastic convex programming [18]	$\mathcal{F}_1 = \{\boldsymbol{x} \in G_1^{\varsigma_1} \subset \mathbb{R}^{m_1} : A\boldsymbol{x} = \boldsymbol{b}\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in G_2^{\varsigma_2} \subset \mathbb{R}^{m_2} : W^{(k)}\boldsymbol{y}^{(k)} = \boldsymbol{h}^{(k)} - T^{(k)}\boldsymbol{x}\}$	$O\left(\sqrt{\varsigma_1 + K\varsigma_2}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((\varsigma_1 + K\varsigma_2)\ \ln(\frac{\mu^0}{\epsilon})\right)$
Stochastic linear programming in a Hilbert space $\mathcal{H}$	$\mathcal{F}_1 = \{\boldsymbol{x} \in \mathcal{G} \subset \mathcal{H} : \boldsymbol{b} - \mathcal{A}\boldsymbol{x} \in \mathbb{R}_+^{m_1}\}$ $\mathcal{F}_2^{(k)}(\boldsymbol{x}) = \{\boldsymbol{y}^{(k)} \in \mathcal{G}^{(k)} \subset \mathcal{H} : \boldsymbol{h}^{(k)} - \mathcal{T}^{(k)}\boldsymbol{x} - \mathcal{W}^{(k)}\boldsymbol{y}^{(k)} \in \mathbb{R}_+^{m_2}\}$	$O\left(\sqrt{m_1 + Km_2}\ \ln(\frac{\mu^0}{\epsilon})\right)$ $O\left((m_1 + Km_2)\ \ln(\frac{\mu^0}{\epsilon})\right)$

that the complexity results in Theorems 6.1 and 6.2 for the short- and long-step algorithms, respectively, are the counterparts of the complexity results in the following theorems for the corresponding two-stage stochastic programs: Theorems 4.1 and 4.2 in [13] for stochastic second-order cone programming, Theorems 4.1 and 4.2 in [15] for stochastic semidefinite programming, Theorems 5 and 6 in [17] for stochastic symmetric cone programming, and Theorems 6.2 and 6.4 in [18] for stochastic convex programming.

In Table 1, we summarize the complexity results for the above classes of stochastic programming problems. We point out that the symbols given in Table 1 were introduced at the end of Section 2. Note that there is a complete matching between the complexity results Theorems 6.1 and 6.2 and their counterparts complexity results summarized in Table 1.

7 Concluding Remarks

In this paper, we have studied the two-stage stochastic linear programming problem in a Hilbert space. We developed and analyzed logarithmic-barrier decomposition interior-point algorithms for solving this class of optimization problems. We proved the convergence and complexity results of the proposed algorithms by presenting some fundamental properties associated with the recourse function. We also described and analyzed short-and long-step versions of the algorithms that follow the primal central trajectory of the first-stage problem.

We found that, for a stochastic linear program in a Hilbert space with m_1 and m_2 inequality constraints in the first- and second-stage problems, respectively, and for K number of realizations, we need at most $O\left(\sqrt{m_1 + Km_2}\, \ln(\mu^0/\epsilon)\right)$ Newton iterations in the short-step algorithm to follow the central path from a starting value of the barrier parameter μ^0 to a terminating value ϵ and we need at most $O\left((m_1 + Km_2)\ \ln(\mu^0/\epsilon)\right)$ Newton iterations in the long-step algorithm for this recentering, where ϵ is the desired accuracy of the final solution. Note that the above complexity results coincide with the best known in the finite-dimensional case. Most notably, these complexity results are independent on the choice of the underlying Hilbert space, which is a conclusion that adds novelty to our algorithms.

The analysis of this paper can be extended to stochastic convex quadratic programming in a Hilbert space and stochastic semidefinite programming with regularized determinants. Nevertheless, from our point of view, it is an important contribution of the future work in this direction to developing decomposition interior-point algorithms for solving stochastic convex programming in Hilbert spaces, in which we might be able to take advantage of the work of Zhao [18] for stochastic convex programming in the finite-dimensional case.

Acknowledgment This work was performed while the author was visiting Center for Applied and Computational Mathematics at Rochester Institute of Technology, NY, USA, and he wishes to thank Akhtar Khan and other colleagues at RIT for their hospitality.

References

1. Renegar, J.: Linear programming, complexity theory and elementary functional analysis. Math. Program. **70**, 279–351 (1995)

2. Renegar, J.: A Mathematical View of Interior-Point Methods in Convex Optimization. SIAM Publications, SIAM, Philadelphia, USA (2001)

3. Faybusovich, L., Moore, J.B.: Long-step path-following algorithm for convex quadratic programming problems in a Hilbert space. J. Optim. Theory Appl. **95**, 615–635 (1997)

4. Faybusovich, L., Moore, J.B.: Infinite-dimensional quadratic optimization: Interior-point methods and control applications. Appl. Math. Optim. **36**, 43–66 (1997)

5. Lim, A.B., Moore, J.B.: A path following algorithm for infinite quadratic programming on a Hilbert space. Discrete Cont. Dyn. Syst. **4**, 653–670 (1998)

6. Santos, P.S., Scheimberg, S.: An outer approximation algorithm for equilibrium problems in Hilbert spaces. Optim. Methods Softw. **30**, 379–390 (2015)

7. Houska, B., Chachuat, B.: Global optimization in Hilbert space. Math. Program., Ser. A, 1–29 (2017)

8. Weiser, M.: Interior point methods in function space. SIAM J. Control Optim. **44**, 1766–1786 (2005)

9. Nesterov, Yu.E., Nemirovskii, A.S.: Interior Point Polynomial Algorithms in Convex Programming. SIAM Publications, Philadelphia, PA (1994)

10. Zhao, G.: A log barrier method with Benders' decomposition for solving two-stage stochastic linear programs. Math. Program. Ser. A **90**, 507–536 (2001)

11. Cho, G.M.: Log-barrier method for two-stage quadratic stochastic programming. Appl. Math. Comput. **164**, 45–69 (2005)

12. Ariyawansa, K.A., Zhu, Y.: A class of volumetric barrier decomposition algorithms for stochastic quadratic programming. Appl. Math. Comput. **186**, 1683–1693 (2007)

13. Alzalg, B.: Decomposition-based interior-point methods for stochastic quadratic second-order cone programming. Appl. Math. Comput. **249**, 1–18 (2014)

14. Alzalg, B.: Volumetric barrier decomposition algorithms for stochastic quadratic second-order cone programming. Appl. Math. Comput. **256**, 494–508 (2015)

15. Mehrotra, S., Özevin, M.G.: Decomposition-based interior point methods for two-stage stochastic semidefinite programming. SIAM J. of Optim. **18**, 206–222 (2007)

16. Ariyawansa, K.A., Zhu, Y.: A class of polynomial volumetric barrier decomposition algorithms for stochastic semidefinite programming. Math. of Comp. **80**, 1639–1661 (2011)

17. Alzalg, B., Ariyawansa, K.A.: Logarithmic barrier decomposition-based interior-point methods for stochastic symmetric programming. J. Math. Anal. Appl. **409**, 973–995 (2014)

18. Zhao, G.: A Lagrangian dual method with self-concordant barrier for multi-stage stochastic convex nonlinear programming. Math. Program. **102**, 1–24 (2005)

19. Alzalg, B., Gafour, A., Alzaleq, L.: Volumetric barrier decomposition algorithms for two-stage stochastic linear semi-infinite programming. Submitted for publication. Available from: <http://www.optimization-online.org/DB_ HTML/2018/12/6978.html>

20. Alzalg, B.: Decomposition interior-point methods based on unital JH-algebras for stochastic conic optimization in spin factors. Unpublished work. Available from: <http://www.optimization-online.org/DB_ HTML/2019/03/7103.html>

21. Kreyszig, E.: Introductory Functional Analysis with Applications. John Wiley and Sons (1978)

22. Han, J., Van, R.B.: Control of diffusions via linear programming. Infanger G, ed. Stochastic Programming: The State of the Art, in Honor of George B. Dantzig (Springer, New York), 329–354 (2011)

23. Ahn, A., Haugh, M.: Linear programming and the control of diffusion processes. INFORMS J Comput. **27**, 646–657 (2015)