Communications in Combinatorics and Optimization
Vol. 9, No. 2 (2024), pp. 353-387
https://doi.org/10.22049/cco.2023.28129.1452

CCO
Commun. Comb. Optim.

Research Article

A path-following algorithm for stochastic quadratically constrained convex quadratic programming in a Hilbert space

Amira Achouak Oulha[†] **and Baha Alzalg**[*]

Department of Mathematics, The University of Jordan, Amman 11942, Jordan
[†]amy9170464@ju.edu.jo
[*]b.alzalg@ju.edu.jo

Received: 30 November 2022; Accepted: 10 April 2023
Published Online: 15 April 2023

Abstract: We propose logarithmic-barrier decomposition-based interior-point algorithms for solving two-stage stochastic quadratically constrained convex quadratic programming problems in a Hilbert space. We prove the polynomial complexity of the proposed algorithms, and show that this complexity is independent on the choice of the Hilbert space, and hence it coincides with the best-known complexity estimates in the finite-dimensional case. We also apply our results on a concrete example from the stochastic control theory.

Keywords: Interior-point methods, Quadratic programming, Stochastic programming, Programming in abstract spaces, Control problems

AMS Subject classification: 90C15, 90C20, 90C48, 90C51, 34K35

1. Introduction

The purpose of this paper is to develop a polynomial-time logarithmic-barrier decomposition-based path-following algorithm for solving two-stage stochastic quadratically constrained convex quadratic programming (SQCCQP) problem in a Hilbert space.
Interior-point methods (also called barrier methods) are one of the most successful class of algorithms to solve deterministic and stochastic optimization problems in both finite and infinite-dimensional settings; see for example [1–3, 5–20]. In the deterministic case, Nesterov and Nemirovskii [16] used the notion of self-concordance to solve different classes of a finite-dimensional optimization problems by interior-point methods. Renegar [17] presented an infinite-dimensional extension of Nesterov and Nemirovskii's work [16]. In addition, Faybusovich and Moore extended

* *Corresponding Author*

path-following interior-point algorithms from infinite-dimensional linear programming to infinite-dimensional convex quadratic programming problems in [12], and more generally to infinite-dimensional quadratically constrained convex quadratic programming in [11].
In the stochastic case, Zhao [20] proposed a logarithmic-barrier algorithm with Benders decomposition for solving two-stage stochastic finite-dimensional linear programming. Cho [10] and Mehrotra and Özevin [13] extended the work of Zhao [20] to two-stage stochastic finite-dimensional convex quadratic programming. Alzalg [4] made a natural step in this direction by extending the work of Zhao [20] to two-stage stochastic infinite-dimensional linear programming.
Given the above brief literature review, the work in this paper can be viewed as an extension of the work of: (i) Faybusovich and Moore [11, 12] from the deterministic case to the stochastic case. (ii) Cho [10] and that of Mehrotra and Özevin [13] from the finite-dimensional Euclidean space to the infinite-dimensional Hilbert space. (iii) Alzalg [4] from the linear case to the quadratic case. Based on the notion of the self-concordance, we prove the polynomial complexity of the proposed algorithms, and also show that this complexity is independent on the choice of the Hilbert space, hence it coincides with the best-known complexity for the finite-dimensional case. Our analysis follows the template in Alzalg [4].
This paper is organized as follows: In Section 3, we define the problem formulation and make some assumptions. In Section 4, we compute the Fréchet derivatives of the recourse function. Self-concordance properties of the recourse function are given in Section 5 with proofs. The proposed logarithmic-barrier path-following interior-point decomposition algorithm is presented in Section 6. Section 7 provides a complexity analysis for the proposed algorithm. In Section 8, we apply the obtained results on a concrete example from stochastic control. Section 9 contains some concluding remarks. In Appendix A, we state some technical lemmas which are required in proving the complexity results of the proposed algorithm. We end this section by providing some notations that will be used in the sequel.

2. Notations

We write $(\mathbf{H}, \langle\cdot,\cdot\rangle)$, or simply $\mathbf{H}$, for a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$, and use $\|\cdot\| \triangleq \sqrt{\langle\cdot,\cdot\rangle}$ for its corresponding norm. If $\mathbf{G}$ is a closed subspace of $\mathbf{H}$, then $x \in \mathbf{H}$ is orthogonal to $\mathbf{G}$, written as $x \perp \mathbf{G}$, if $\langle x, y\rangle = 0$ for all $y \in \mathbf{G}$. The set $\mathbf{G}^\perp$ is the orthogonal complement of $\mathbf{G}$ and is defined as

$$\mathbf{G}^\perp \triangleq \{x \in \mathbf{H} : x \perp \mathbf{G}\}.$$

Let $f : \mathbf{H} \to \mathbb{R}$ be a smooth function and $x \in \mathbf{H}$, then the gradient $\nabla_x f(x)$ is uniquely determined as

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

where $\mathcal{D}_x f(x)(\xi)$ stands for the first Fréchet derivative of f at the point x evaluated on ξ. The second derivative is given by

$$\mathcal{D}^2_{xx} f(x)(\xi, \zeta) \triangleq \mathcal{D}_x(\mathcal{D}_x f(x)(\xi))(\zeta),$$

where $\mathcal{D}^2_{xx} f(x)(\xi)$ stands for the second Fréchet derivative of f at the point x evaluated on (ξ, ζ). We can also define higher Fréchet derivatives in a similar way.
Let $y, a_1, a_2, \ldots, a_m \in \mathbf{H}$. Throughout this paper, we denote by "$\mathcal{A}y$" the vector in $\mathbb{R}^m$ whose i^{th}-entry is the scalar $\langle a_i, y\rangle$, for $i = 1, 2, \ldots, m$. For $z \in \mathbb{R}^m$, we also denote by "$\mathcal{A}^\dagger z$" the element $\sum_{i=1}^m z_i a_i$ in $\mathbf{H}$. Note that $\mathcal{A}$ maps the Hilbert space $\mathbf{H}$ onto $\mathbb{R}^m$, while $\mathcal{A}^\dagger$ maps $\mathbb{R}^m$ into the Hilbert space $\mathbf{H}$. For $y \in \mathbf{H}$ and $h = h(x) \in \mathbb{R}^m$ with $x \in \mathbf{H}$, we denote by "$\mathcal{J}_x[h]y$" the vector in $\mathbb{R}^m$ whose ith entry is $\langle \nabla_x h_i(x), y\rangle$ for $i = 1, 2, \ldots, m$. For $z \in \mathbb{R}^m$, we also denote by "$\mathcal{J}_x^\dagger[h]z$" the element $\sum_{i=1}^m z_i \nabla_x h_i(x)$ in $\mathbf{H}$. If $x \in \mathbf{H}$ and $\mathscr{P}_1, \mathscr{P}_2, \ldots, \mathscr{P}_m$ are self-adjoint nonnegative definite bounded operators on $\mathbf{H}$, we denote by $\mathcal{P}[x]$ the vector in $\mathbb{R}^m$ whose ith entry is $\mathscr{P}_i x$ for $i = 1, 2, \ldots, m$. For $x, y \in \mathbf{H}$, by "$\mathcal{P}[x]y$" we mean the vector in $\mathbb{R}^m$ whose ith entry is $\langle \mathscr{P}_i x, y\rangle$ for $i = 1, 2, \ldots, m$. For $z \in \mathbb{R}^m$, by "$\mathcal{P}^\dagger[x]z$" we mean the element $\sum_{i=1}^m z_i \mathscr{P}_i x$ in $\mathbf{H}$.
In summary, we have

$$\mathcal{A}y \triangleq \begin{bmatrix} \langle a_1, y\rangle \\ \langle a_2, y\rangle \\ \vdots \\ \langle a_m, y\rangle \end{bmatrix} \in \mathbb{R}^m, \;\; \mathcal{J}_x[h]y \triangleq \begin{bmatrix} \langle \nabla_x h_1(x), y\rangle \\ \langle \nabla_x h_2(x), y\rangle \\ \vdots \\ \langle \nabla_x h_m(x), y\rangle \end{bmatrix} \in \mathbb{R}^m, \;\; \mathcal{P}[x]y \triangleq \begin{bmatrix} \langle \mathscr{P}_1 x, y\rangle \\ \langle \mathscr{P}_2 x, y\rangle \\ \vdots \\ \langle \mathscr{P}_m x, y\rangle \end{bmatrix} \in \mathbb{R}^m,$$
$$\mathcal{A}^\dagger z \triangleq \sum_{i=1}^m z_i a_i \in \mathbf{H}, \quad \mathcal{J}_x^\dagger[h]z \triangleq \sum_{i=1}^m z_i \nabla_x h_i(x) \in \mathbf{H}, \quad \mathcal{P}^\dagger[x]z \triangleq \sum_{i=1}^m z_i \mathscr{P}_i x \in \mathbf{H}.$$

The above operators enjoy pointwise operations. For example, but not limited to, it is possible to define

$$(\mathcal{A} + \mathcal{J}_x[h] + \mathcal{P}[x])y \triangleq \mathcal{A}y + \mathcal{J}_x[h]y + \mathcal{P}[x]y, \tag{1}$$

and

$$(\mathcal{A}^\dagger + \mathcal{J}_x^\dagger[h] + \mathcal{P}^\dagger[x])z \triangleq \mathcal{A}^\dagger z + \mathcal{J}_x^\dagger[h]z + \mathcal{P}^\dagger[x]z,$$

and so on. Note also that

$$\left\langle \mathcal{A}^\dagger z, y \right\rangle = \left\langle \sum_{i=1}^m z_i a_i, y \right\rangle = \sum_{i=1}^m \left(z_i \langle a_i, y\rangle\right) = z^\top \mathcal{A}y. \tag{2}$$

Similarly, we also have

$$\left\langle \mathcal{P}^\dagger[x]z, y \right\rangle = z^\top \mathcal{P}[x]y \;\text{ and }\; \left\langle \mathcal{J}_x^\dagger[h]z, y \right\rangle = z^\top \mathcal{J}_x[h]y. \tag{3}$$

Throughout the paper, we use $\mathbb{R}_{++}$ to denote the set of all positive real numbers. We write e for a vector with all entries equal to one. The dimension of e will be clear from the context. For any vector $z \in \mathbb{R}^m$, we define $Z \triangleq \text{Diag}\,(z_1, z_2, \ldots, z_m)$. That is, Z denotes the $m \times m$ diagonal matrix whose diagonal entries are $z_1, z_2, \ldots, z_m$.

3. Problem formulation and assumptions

In this section, we write a formulation for the two-stage stochastic quadratically constrained convex quadratic problem in a Hilbert space, followed by a logarithmic-barrier formulation based on our settings. Then we make some assumptions.

Let $\mathbf{G}$ be a close subspace of a real Hilbert space $(\mathbf{H}, \langle \cdot, \cdot \rangle)$. Let also $\mathscr{P}_i : \mathbf{H} \to \mathbf{H}, i = 0, 1, \ldots, m_1$, and $\mathscr{Q}_j(\omega) : \mathbf{H} \to \mathbf{H}, j = 0, 1, \ldots, m_2$, be self-adjoint nonnegative definite bounded operators on $\mathbf{H}$. We consider the two-stage SQCCQP problem in standard form:

$$\begin{array}{ll} \min & \frac{1}{2}\langle \mathscr{P}_0 x, x\rangle + \langle a_0, x\rangle + \mathbb{E}[\rho(x,\omega)] \\ \text{s.t.} & \frac{1}{2}\langle \mathscr{P}_i x, x\rangle + \langle a_i, x\rangle \leq b_i,\ i = 1, 2, \ldots, m_1, \\ & x \in \mathbf{G}, \end{array} \tag{4}$$

where x is the first-stage decision variable, $a_0, a_1, \ldots, a_{m_1} \in \mathbf{H}, b_0 = 0, b_1, b_2, \ldots, b_{m_1} \in \mathbb{R}$, and $\rho(x, \omega)$ is the maximum value of the problem:

$$\begin{array}{ll} \min & \frac{1}{2}\langle \mathscr{Q}_0(\omega) y, y\rangle + \langle w_0(\omega), y\rangle \\ \text{s.t.} & \frac{1}{2}\langle \mathscr{Q}_j(\omega) y, y\rangle + \langle w_j(\omega), y\rangle \leq h_j(\omega) - \langle t_j(\omega), x\rangle,\ j = 1, 2, \ldots, m_2, \\ & y \in \mathbf{G}, \end{array} \tag{5}$$

where y is the second-stage variable, $w_0(\omega), w_1(\omega), \ldots, w_{m_2}(\omega), t_1(\omega), \ldots, t_{m_2}(\omega) \in \mathbf{H}$, $t_0(\omega) = 0, h_0(\omega) = 0$, and $h_1(\omega), h_2(\omega), \ldots, h_{m_2}(\omega) \in \mathbb{R}$. The realizations of the random data $w_j(\omega), t_j(\omega), h_j(\omega)$ and $\mathscr{Q}_j(\omega)$, for $j = 0, 1, \ldots, m_2$, depend on an underling outcome ω in an event space Ω with a known probability function $\mathbb{P}$.

Let $p_i(x) \triangleq \frac{1}{2}\langle \mathscr{P}_i x, x\rangle + \langle a_i, x\rangle - b_i$ for $i = 0, 1, \ldots, m_1$, and $q_j(y, x, \omega) \triangleq \frac{1}{2}\langle \mathscr{Q}_j(\omega) y, y\rangle + \langle w_j(\omega), y\rangle + \langle t_j(\omega), x\rangle - h_j(\omega)$ for $j = 0, 1, \ldots, m_2$. Then (4) and (5) are written respectively as

$$\begin{array}{ll} \min & p_0(x) + \mathbb{E}[\rho(x,\omega)] \\ \text{s.t.} & p_i(x) \leq 0,\ i = 1, 2, \ldots, m_1, \\ & x \in \mathbf{G}, \end{array} \tag{6}$$

and

$$\begin{array}{ll} \min & q_0(y, x, \omega) \\ \text{s.t.} & q_j(y, x, \omega) \leq 0,\ j = 1, 2, \ldots, m_2, \\ & y \in \mathbf{G}. \end{array}$$

We examine (4) and (5) when Ω is discrete and finite. Let $\{(t_j^{(k)}, w_0^{(k)}, w_j^{(k)}, h_j^{(k)}, \mathscr{Q}_0^{(k)}, \mathscr{Q}_j^{(k)}) : 1 \leq k \leq K\}$ be the set of K possible realizations of random variables $(t_j(\omega), w_0(\omega), w_j(\omega), h_j(\omega), \mathscr{Q}_0(\omega), \mathscr{Q}_j(\omega))$. Let also

$$\pi_k \triangleq \mathbb{P}\left(\left(t_{1\leq j\leq m_2}(\omega), w_{1\leq j\leq m_2}(\omega), h_{1\leq j\leq m_2}(\omega), Q(\omega)\right) = \left(t^{(k)}_{1\leq j\leq m_2}, w^{(k)}_{1\leq j\leq m_2}, h^{(k)}_{1\leq j\leq m_2}, \mathscr{Q}^{(k)}\right)\right)$$

be the associated probability for $K = 1, 2, \ldots, K$. Then (4) and (5) can be written as

$$\begin{array}{ll} \min & \eta(x) \triangleq p_0(x) + \sum_{k=1}^{K} \rho^{(k)}(x) \\ \text{s.t.} & p_i(x) \leq 0,\ i = 1, 2, \ldots, m_1, \\ & x \in \mathbf{G}, \end{array} \tag{7}$$

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

$$\begin{array}{ll} \min & q_0^{(k)}(y, x) \triangleq \frac{1}{2}\langle \mathcal{Q}_0^{(k)} y, y\rangle + \langle w_0^{(k)}, y\rangle \\ \text{s.t.} & q_j^{(k)}(y, x) \triangleq \frac{1}{2}\langle \mathcal{Q}_j^{(k)} y, y\rangle + \langle w_j^{(k)}, y\rangle - h_j^{(k)} + \langle t_j^{(k)}, x\rangle \leq 0,\ j = 1, 2, \ldots, m_2, \\ & y \in \mathbf{G}. \end{array} \tag{8}$$

Along with Problem (8), we consider the (dual) problem

$$\begin{array}{ll} \max & q_0^{(k)}(y, x) + \sum_{j=1}^{m_2} \lambda_j^{(k)} q_j^{(k)}(y, x) \\ \text{s.t.} & \nabla q_0^{(k)}(y, x) + \sum_{j=1}^{m_2} \lambda_j^{(k)} \nabla q_j^{(k)}(y, x) \in \mathbf{G}^{\perp}, \\ & \lambda^{(k)} \geq 0,\ y \in \mathbf{G}, \end{array} \tag{9}$$

where $\lambda^{(k)} \in \mathbb{R}^{m_2}$ is the second-stage dual multiplier.
The logarithmic-barrier problem associated with (7) and (8) is

$$\begin{array}{ll} \min & \eta(x, \mu) \triangleq p_0(x) - \mu \sum_{i=1}^{m_1} \ln s_i + \sum_{k=1}^{K} \rho^{(k}(x, \mu) \\ \text{s.t.} & p_i(x) + s_i = 0,\ i = 1, 2, \ldots, m_1 \\ & s > 0,\ x \in \mathbf{G}, \end{array} \tag{10}$$

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

$$\begin{array}{ll} \min & \rho^{(k)}(x, \mu) \triangleq q_0^{(k)}(y, x) - \mu \sum_{j=1}^{m_2} \ln z_j^{(k)} \\ \text{s.t.} & q_j^{(k)}(y, x) + z_j^{(k)} = 0,\ j = 1, 2, \ldots, m_2, \\ & z^{(k)} > 0,\ y \in \mathbf{G}, \end{array} \tag{11}$$

where $z^{(k)} \triangleq -q^{(k)}(y, x)$ for $k = 1, 2, \ldots, K$. Note that if for some k, Problem (10) is infeasible, we define $\sum_{k=1}^{K} \rho^{(k)}(x, \mu) \triangleq \infty$. The barrier problem associated with the (dual) problem (9) is the problem

$$
\begin{aligned}
\max \quad & q_0^{(k)}(y,x) + \sum_{j=1}^{m_2} \lambda_j^{(k)} q_j^{(k)}(y,x) + \mu \sum_{j=1}^{m_2} \ln \lambda_j^{(k)} \\
\text{s.t.} \quad & \nabla q_0^{(k)}(y,x) + \sum_{j=1}^{m_2} \lambda_j^{(k)} \nabla q_j^{(k)}(y,x) \in \mathbf{G}^{\perp}, \\
& \lambda^{(k)} > 0, \ y \in \mathbf{G}, \ j = 1,2,\ldots,m_2.
\end{aligned} \tag{12}
$$

In light of Propositions 2.6 and 2.7 (see also Corollary 2.4) in [11], the points $y, z^{(k)}$ and $\lambda^{(k)}$ are the optimal solutions of Problems (11) and (12) if and only if they satisfy the following optimality conditions:

$$
\begin{array}{l}
\frac{1}{2}\langle \mathcal{Q}_j^{(k)} y, y\rangle + \langle w_j^{(k)}, y\rangle + z_j^{(k)} = h_j^{(k)} - \langle t_j^{(k)}, x\rangle, \ j = 1,2,\ldots,m_2, \\
\left\langle y, \ w_0^{(k)} + \mathcal{Q}_0^{(k)} y + \sum_{j=1}^{m_2} \left(\lambda_j^{(k)}\left(w_j^{(k)} + \mathcal{Q}_j^{(k)} y\right)\right)\right\rangle = 0, \\
Z^{(k)}\lambda^{(k)} = \mu e, \\
z^{(k)}, \lambda^{(k)} > 0,
\end{array} \tag{13}
$$

where the second equation follows by using the fact that

$$
y \in \mathbf{G}, \quad \text{while} \quad w_0^{(k)} + \mathcal{Q}_0^{(k)} y + \sum_{j=1}^{m_2} \left(\lambda_j^{(k)} w_j^{(k)} + \lambda_j^{(k)} \mathcal{Q}_j^{(k)} y\right) = \nabla q_0^{(k)}(y,x) + \sum_{j=1}^{m_2} \lambda_j^{(k)} \nabla q_j^{(k)}(y,x) \in \mathbf{G}^{\perp}.
$$

Problems (10) and (11) can be equivalently written as a deterministic infinite-dimensional optimization problem

$$
\begin{aligned}
\min \quad & p_0(x) + \sum_{k=1}^{K} q_0^{(k)}(y,x) - \mu \left(\sum_{i=1}^{m_2} \ln s_i + \sum_{k=1}^{K} \sum_{j=1}^{m_2} \ln z_j^{(k)} \right) \\
\text{s.t.} \quad & p_i(x) + s_i = 0, \ i = 1,2,\ldots,m_1, \\
& q_j^{(k)}(y,x) + z_j^{(k)} = 0, \ j = 1,2,\ldots,m_2, \ k = 1,2,\ldots,K, \\
& x, y \in \mathbf{G}, \\
& s, z^{(k)} > 0, \ k = 1,2,\ldots,K.
\end{aligned} \tag{14}
$$

We define the following feasibility sets:

$$
\begin{array}{lcl}
\mathcal{F}_1 & \triangleq & \{x \in \mathbf{G} : \ s = -p(x) > 0\}; \\
\mathcal{F}_2^{(k)}(x) & \triangleq & \left\{y \in \mathbf{G} : z^{(k)} = -q^{(k)}(y,x) > 0\right\}, \text{ for } k = 1,2,\ldots,K; \\
\mathcal{F}_2^{(k)} & \triangleq & \left\{x \in \mathbf{G} : \mathcal{F}_2^{(k)}(x) \neq \emptyset\right\}, \text{ for } k = 1,2,\ldots,K; \\
\mathcal{F}_2 & \triangleq & \cap_{k=1}^{K} \mathcal{F}_2^{(k)}; \\
\mathcal{F}_0 & \triangleq & \mathcal{F}_1 \cap \mathcal{F}_2; \\
\mathcal{F} & \triangleq & \left\{(x,r) \times \left(y, \lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(K)}\right) : -p(x) > 0, \ -q^{(k)}(y,x) > 0, \right. \\
& & \left\langle y, \nabla q_0^{(k)}(y,x) + \sum_{j=1}^{m_2} \lambda_j^{(k)} \nabla q_j^{(k)}(y,x) \right\rangle = 0, \ \lambda^{(k)} > 0, \ 1 \le k \le K, \\
& & \left. \left\langle x, \ \nabla p_0(x) + \sum_{i=1}^{m_1} r_i \nabla p_i(x) + \sum_{k=1}^{K} \sum_{j=1}^{m_2} \lambda_j^{(k)} t_j^{(k)} \right\rangle = 0, \ r > 0 \right\}.
\end{array}
$$

Here $r \in \mathbb{R}^{m_1}$ is the first-stage dual multiplier.
Now, we are ready to make two assumptions:

Assumption 1. *The elements $a_1, a_2, \ldots, a_{m_1}$ are linearly independent in $\boldsymbol{H}$, and for $k = 1, 2, \ldots, K$, the elements $t_1^{(k)}, t_2^{(k)}, \ldots, t_{m_2}^{(k)}$ are linearly independent in $\boldsymbol{H}$, and the elements $w_1^{(k)}, w_2^{(k)}, \ldots, w_{m_2}^{(k)}$ are linearly independent in $\boldsymbol{H}$.*

Assumption 2. *The feasibility set $\mathcal{F}$ is nonempty.*

Assumption 1 is important to ensure the operator invertibility. Assumption 2 means that Problem (14) and its dual problem have strictly feasible solutions, which ensures that strong duality holds for first- and second-stage SQCCQP problems. This indicates that each of the optimization problems (10)-(14) has a unique solution. It is worth noting that for each given $\mu > 0$, $\sum_{k=1}^{K} \rho^{(k)}(x, \mu) < \infty$ if and only if $x \in \mathcal{F}_2$. Therefore, the optimal solutions of Problems (10) and (11) and that of Problem (14) have a relationship that is described in the following remark.

Remark 1. The point $(x(\mu), y(\mu), s, z^{(1)}, z^{(2)}, \ldots, z^{(K)})$ is the optimal solution of (14) if and only if $(x(\mu), s)$ is the optimal solution of (10) and $(y(\mu), z^{(1)}, z^{(2)}, \ldots, z^{(K)})$ is the optimal solution for (11) for given μ and $x \triangleq x(\mu)$.

The next section shows such a relationship and computes the Fréchet derivatives $\mathcal{D}_x\eta(x, \mu)$ and $\mathcal{D}^2_{xx}\eta(x, \mu)$ which are required to calculate the Newton direction.

4. Derivatives of the recourse function

In this section, we compute the Fréchet derivatives $\mathcal{D}_x\eta(x, \mu)$ and $\mathcal{D}^2_{xx}\eta(x, \mu)$. In order to do that, we first need to compute the Fréchet derivatives of the recourse function $\rho^{(k)}(x, \mu)$ with respect to x.
Let $(y, z^{(k)}, \lambda^{(k)}) \triangleq (y(x, \mu), z^{(k)}(x, \mu), \lambda^{(k)}(x, \mu))$. Differentiate (13) with respect to x, we get

$$\begin{array}{rcll} \left\langle \mathcal{Q}_j^{(k)} y, \mathcal{D}_x \langle y, \xi \rangle \right\rangle + \left\langle w_j^{(k)}, \mathcal{D}_x \langle y, \xi \rangle \right\rangle + \left\langle \nabla_x z_j^{(k)}, \xi \right\rangle & = & -\left\langle t_j^{(k)}, \xi \right\rangle, & j = 1, 2, \ldots, m_2, \\ \left\langle y,\ \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y, \xi \rangle + \ \sum_{j=1}^{m_2} \left(\left\langle \nabla_x \lambda_j^{(k)}, \xi \right\rangle \left(w_j^{(k)} + \mathcal{Q}_j^{(k)} y \right) + \lambda_j^{(k)} \mathcal{Q}_j^{(k)} \mathcal{D}_x \langle y, \xi \rangle \right) \right\rangle & = & 0, & \\ \left\langle \nabla_x \lambda_j^{(k)}, \xi \right\rangle z_j^{(k)} + \left\langle \nabla_x z_j^{(k)}, \xi \right\rangle \lambda_j^{(k)} & = & 0, & j = 1, 2, \ldots, m_2, \end{array} \tag{15}$$

for any $\xi \in \mathbf{H}$, where the second equation was obtained in view of the fact that

$$\left\langle \mathcal{D}_x \langle y, \xi \rangle,\ \left(w_0^{(k)} + \mathcal{Q}_0^{(k)} y + \sum_{j=1}^{m_2} \left(\lambda_j^{(k)} \left(w_j^{(k)} + \mathcal{Q}_j^{(k)} y \right) \right) \right) \right\rangle = 0$$

due to the orthogonality relation.
Following our notations introduced in Subsection 2, System (15) can be written more compactly as

$$\begin{array}{rcl} \mathcal{Q}^{(k)}[y] \mathcal{D}_x \langle y, \xi \rangle + \mathcal{W}^{(k)} \mathcal{D}_x \langle y, \xi \rangle + \mathcal{J}_x \left[z^{(k)} \right] \xi & = & -\mathcal{T}^{(k)} \xi, \\ \left\langle y,\ \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y, \xi \rangle + \left(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y] \right) \mathcal{J}_x \left[\lambda^{(k)} \right] \xi + \mathcal{Q}^{(k)\dagger} [\mathcal{D}_x \langle y, \xi \rangle] \lambda^{(k)} \right\rangle & = & 0, \\ Z^{(k)} \mathcal{J}_x \left[\lambda^{(k)} \right] \xi + \Lambda^{(k)} \mathcal{J}_x \left[z^{(k)} \right] \xi & = & 0. \end{array} \tag{16}$$

Solving System (16), we get

$$\begin{array}{rcl} \mathcal{D}_x\langle y,\xi\rangle &=& -\mathscr{R}^{(k)-1}\left(Q^{(k)\dagger}[y]+\mathcal{W}^{(k)\dagger}\right)L^{(k)^2}\mathcal{T}^{(k)}\xi,\\ \mathcal{J}_x\left[z^{(k}\right]\xi &=& -L^{(k)-1}N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi,\\ \mathcal{J}_x\left[\lambda^{(k)}\right]\xi &=& L^{(k)}N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi, \end{array} \tag{17}$$

where

$$\begin{array}{lllll} L^{(k)} &\triangleq& L^{(k)}(\mu,\xi) &\triangleq& \left(Z^{(k)-1}\Lambda^{(k)}\right)^{\frac{1}{2}},\\ N^{(k)} &\triangleq& N^{(k)}(x,\mu) &\triangleq& I-L^{(k)}\left(Q^{(k)}[y]+\mathcal{W}^{(k)}\right)\mathscr{R}^{(k)-1}\left(Q^{(k)\dagger}[y]+\mathcal{W}^{(k)\dagger}\right)L^{(k)},\\ \mathscr{R}^{(k)}\xi &\triangleq& \mathscr{R}^{(k)}(x,\mu)\xi &\triangleq& \left(Q^{(k)\dagger}[y]+\mathcal{W}^{(k)\dagger}\right)L^{(k)2}\left(Q^{(k)}[y]+\mathcal{W}^{(k)}\right)\xi+\mathscr{Q}_0^{(k)}\xi+Q^{(k)\dagger}[\xi]\lambda^{(k)}, \end{array} \tag{18}$$

for $\xi\in\mathbf{H}$. Note that, based on Assumption 1, $\mathscr{R}^{(k)}\cdot$ is an invertible operator from the Hilbert space $\mathbf{H}$ into itself, hence the operator $\mathscr{R}^{(k)-1}\cdot$ is well-defined.
Now we are ready to compute the Fréchet derivatives of $\rho(x,\mu)$. Note that

$$\begin{array}{rcl} \displaystyle\sum_{i=1}^{m_2}\lambda_j^{(k)}\left(h_j^{(k)}-\langle t_j^{(k)},x\rangle\right) &=& \displaystyle\sum_{i=1}^{m_2}\lambda_j^{(k)}\left(\frac{1}{2}\langle\mathscr{Q}_j^{(k)}y,y\rangle+\langle w_j^{(k)},y\rangle+z_j^{(k)}\right)\\ &=& \displaystyle\tfrac{1}{2}\sum_{j=1}^{m_2}\lambda_j^{(k)}\langle\mathscr{Q}_j^{(k)}y,y\rangle+\sum_{j=1}^{m_2}\lambda_j^{(k)}\langle w_j^{(k)},y\rangle+\sum_{j=1}^{m_2}z_j^{(k)}\lambda_j^{(k)}\\ &=& \displaystyle\left\langle w_0^{(k)}+\mathscr{Q}_0^{(k)}y+\sum_{j=1}^{m_2}\left(\lambda_j^{(k)}w_j^{(k)}+\lambda_j^{(k)}\mathscr{Q}_j^{(k)}y\right),y\right\rangle\\ && \displaystyle-\left\langle w_0^{(k)},y\right\rangle-\left\langle\mathscr{Q}_0^{(k)}y,y\right\rangle+\sum_{j=1}^{m_2}z_j^{(k)}\lambda_j^{(k)}\\ && \displaystyle-\sum_{j=1}^{m_2}\lambda_j^{(k)}\left\langle\mathscr{Q}_j^{(k)}y,y\right\rangle+\frac{1}{2}\sum_{j=1}^{m_2}\lambda_j^{(k)}\left\langle\mathscr{Q}_j^{(k)}y,y\right\rangle\\ &=& \displaystyle\sum_{j=1}^{m_2}z_j^{(k)}\lambda_j^{(k)}-\left\langle w_0^{(k)},y\right\rangle-\left\langle\mathscr{Q}_0^{(k)}y,y\right\rangle-\frac{1}{2}\sum_{j=1}^{m_2}\lambda_j^{(k)}\left\langle\mathscr{Q}_j^{(k)}y,y\right\rangle\\ &=& \displaystyle\mu m_2-\left\langle w_0^{(k)},y\right\rangle-\left\langle\mathscr{Q}_0^{(k)}y,y\right\rangle-\tfrac{1}{2}\sum_{j=1}^{m_2}\lambda_j^{(k)}\left\langle\mathscr{Q}_j^{(k)}y,y\right\rangle. \end{array}$$

Using the definition of $\rho^{(k)}(\cdot,\cdot)$ in (14), we have

$$\begin{array}{l} \displaystyle\sum_{i=1}^{m_2}\lambda_j^{(k)}\left(\langle t_j^{(k)},x\rangle-h_j^{(k)}\right)-\frac{1}{2}\left\langle\mathscr{Q}_0^{(k)}y,y\right\rangle-\frac{1}{2}\sum_{j=1}^{m_2}\lambda_j^{(k)}\left\langle\mathscr{Q}_j^{(k)}y,y\right\rangle+\mu\sum_{j=1}^{m_2}\ln\lambda_j^{(k)}\\ \displaystyle=\left\langle w_0^{(k)},y\right\rangle+\tfrac{1}{2}\left\langle\mathscr{Q}_0^{(k)}y,y\right\rangle-\mu m_2+\mu\sum_{j=1}^{m_2}\ln\lambda_j^{(k)}\\ \displaystyle=\rho^{(k)}(x,\mu)+\mu\sum_{j=1}^{m_2}\ln z_j^{(k)}-\mu m_2+\mu\sum_{j=1}^{m_2}\ln\lambda_j^{(k)}\\ \displaystyle=\rho^{(k)}(x,\mu)-\mu m_2+\mu\sum_{j=1}^{m_2}\ln\left(z_j^{(k)}\lambda_j^{(k)}\right)\\ =\rho^{(k)}(x,\mu)-\mu m_2+\mu m_2\ln\ \mu=\rho^{(k)}(x,\mu)-\mu m_2(1-ln\ \mu). \end{array}$$

It follows that

$$\rho^{(k)}(x,\mu) = \sum_{i=1}^{m_2} \lambda_j^{(k)} \left(\langle t_j^{(k)}, x\rangle - h_j^{(k)}\right) - \frac{1}{2}\left\langle \mathcal{Q}_0^{(k)} y, y\right\rangle - \frac{1}{2}\sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle + \mu \sum_{j=1}^{m_2} \ln \lambda_j^{(k)} + \mu m_2 (1 - \ln\ \mu).$$

Using (13) and (15), we get

$$\begin{aligned}
\mathcal{D}_x\rho^{(k)}(x,\mu) &= \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left(\left\langle t_j^{(k)}, x\right\rangle - h_j^{(k)}\right) + \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) \\
&\quad - \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle - \tfrac{1}{2}\sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle \\
&\quad - \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle + \mu \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \lambda_j^{(k)-1} \\
&= \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) - \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left(\frac{1}{2}\left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle + \left\langle w_j^{(k)}, y\right\rangle + z_j^{(k)}\right) - \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle \\
&\quad - \tfrac{1}{2}\sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle - \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle + \mu \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \lambda_j^{(k)-1} \\
&= \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) - \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left(\frac{1}{2}\left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle + \left\langle w_j^{(k)}, y\right\rangle\right) - \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle \\
&\quad - \tfrac{1}{2}\sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle - \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle + \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left(\mu\lambda_j^{(k)-1} - z_j^{(k)}\right) \\
&= \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) - \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left\langle w_j^{(k)}, y\right\rangle - \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle - \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle \\
&\quad - \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x y, y\right\rangle \\
&= \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) - \sum_{j=1}^{m_2} \mathcal{D}_x\lambda_j^{(k)} \left(\left\langle w_j^{(k)}, y\right\rangle + \left\langle \mathcal{Q}_j^{(k)} y, y\right\rangle\right) \\
&\quad - \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle - \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x y, y\right\rangle \\
&= \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) + \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle + \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle - \left\langle \mathcal{Q}_0^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle \\
&\quad - \sum_{j=1}^{m_2} \lambda_j^{(k)} \left\langle \mathcal{Q}_j^{(k)} \mathcal{D}_x \langle y,\xi\rangle, y\right\rangle \\
&= \sum_{j=1}^{m_2} \lambda_j^{(k)} \left(\mathcal{D}_x \left\langle t_j^{(k)}, x\right\rangle\right) = \sum_{j=1}^{m_2} \lambda_j^{(k)} t_j^{(k)}.
\end{aligned}$$

Then the first Fréchet derivative of $\rho^{(k)}$ is $\mathcal{D}_x\rho^{(k)}(x,\mu) = \mathcal{T}^{(k)\dagger}\lambda^{(k)}$. Thus, for any $\xi \in \mathbf{H}$, the gradient of $\rho^{(k)}$ is uniquely determined as

$$\langle \nabla_x \rho(x,\mu), \xi\rangle = \left\langle \mathcal{T}^{(k)\dagger}\lambda^{(k)}, \xi\right\rangle = \lambda^{(k)^{\mathsf{T}}} \left(\mathcal{T}^{(k)}\xi\right).$$

The second Fréchet derivative of $\rho^{(k)}(\cdot,\cdot)$ is

$$\begin{aligned}\mathcal{D}^2_{xx}\rho^{(k)}(x,\mu)(\xi,\zeta) &= \sum_{j=1}^{m_2}\left(\mathcal{D}_x(\lambda_j^{(k)})(\zeta)\right)\left\langle t_j^{(k)},\xi\right\rangle\\ &= \sum_{j=1}^{m_2}\left\langle \nabla_x\lambda_j^{(k)},\zeta\right\rangle\left\langle t_j^{(k)},\xi\right\rangle\\ &= \left(\mathcal{J}_x\left[\lambda^{(k)}\right]\zeta\right)^{\mathsf{T}}\left(\mathcal{T}^{(k)}\xi\right),\end{aligned}$$

for any $\xi,\zeta\in\mathbf{H}$.
To sum up, the first and second-order Fréchet derivatives of $\rho^{(k)}(\cdot,\cdot)$ are

$$\begin{aligned}\mathcal{D}_x\rho^{(k)}(x,\mu)(\xi) &= \left\langle \mathcal{T}^{(k)\dagger}\lambda^{(k)},\xi\right\rangle &&= \left(\mathcal{T}^{(k)}\xi\right)^{\mathsf{T}}\lambda^{(k)},\\ \mathcal{D}^2_{xx}\rho^{(k)}(x,\mu)(\xi,\zeta) &= \left\langle \mathcal{T}^{(k)\dagger}\mathcal{J}_x\left[\lambda^{(k)}\right]\zeta,\xi\right\rangle &&= \left(\mathcal{T}^{(k)}\xi\right)^{\mathsf{T}}\mathcal{J}_x\left[\lambda^{(k)}\right]\zeta.\end{aligned}\tag{19}$$

Now we are ready to compute the Fréchet derivatives of $\eta(x,\mu)$. By differentiating $p_i(x)+s_i=0$ with respect to x, we get

$$\mathcal{D}_x p_i(x) = \mathcal{D}_x(-s_i(x))(\xi) = \langle -\nabla_x s_i,\xi\rangle = \langle \mathscr{P}_i x + a_i,\xi\rangle\,,$$

for any $\xi\in\mathbf{H}$, as $p_i(x) = -s_i(x) = \frac{1}{2}\langle \mathscr{P}_i x, x\rangle + \langle a_i, x\rangle - b_i$, for $i=1,2,\ldots,m_1$. As a result, we have

$$\mathcal{J}_x[p] = \mathcal{P}[x]\xi + \mathcal{A}\xi,\ \text{ or equivalently } \mathcal{J}_x[s] = -(\mathcal{P}[x]+\mathcal{A})\xi,$$

for any $\xi\in\mathbf{H}$. Using (19), it follows that

$$\begin{aligned}\mathcal{D}_x\eta(x,\mu) &= a_0 + \mathscr{P}_0 x - \mu\sum_{i=1}^{m_1}\frac{\nabla_x s_i(x)}{s_i(x)} - \sum_{k=1}^{K}\mathcal{D}_x\rho^{(k)}(x,\mu)\\ &= a_0 + \mathscr{P}_0 x - \mu\sum_{i=1}^{m_1}\frac{\nabla_x s_i(x)}{s_i(x)} - \sum_{k=1}^{K}\mathcal{T}^{(k)\dagger}\lambda^{(k)}\\ &= a_0 + \mathscr{P}_0 x + \mu\sum_{i=1}^{m_1}\frac{a_i+\mathscr{P}_i x}{s_i(x)} - \sum_{k=1}^{K}\mathcal{T}^{(k)\dagger}\lambda^{(k)}.\end{aligned}\tag{20}$$

Consequently, using (17), we also have

$$\begin{aligned}\mathcal{D}^2_{xx}\eta(x,\mu)(\xi,\zeta) &= \left\langle\left(\mathscr{P}_0 + \mu\sum_{i=1}^{m_1}\frac{\mathscr{P}_i}{s_i(x)}\right)\xi,\zeta\right\rangle + \mu\sum_{i=1}^{m_1}\frac{\langle\nabla_x s_i(x),\xi\rangle\langle\nabla_x s_i(x),\zeta\rangle}{s_i^2(x)}\\ &\quad - \sum_{k=1}^{K}\left\langle \mathcal{T}^{(k)\dagger}\mathcal{J}_x\left[\lambda^{(k)}\right]\zeta,\xi\right\rangle\\ &= \left\langle\left(\mathscr{P}_0 + \mu\sum_{i=1}^{m_1}\frac{\mathscr{P}_i}{s_i(x)}\right)\xi,\zeta\right\rangle + \mu\sum_{i=1}^{m_1}\frac{\langle a_i+\mathscr{P}_i x,\xi\rangle\langle a_i+\mathscr{P}_i x,\zeta\rangle}{s_i^2(x)}\\ &\quad - \sum_{k=1}^{K}\left\langle \mathcal{T}^{(k)\dagger}\mathcal{J}_x\left[\lambda^{(k)}\right]\zeta,\xi\right\rangle,\end{aligned}\tag{21}$$

for any $\xi, \zeta \in \mathbf{H}$.
Note that

$$\sum_{i=1}^{m_1} \frac{a_i + \mathscr{P}_i x}{s_i(x)} = \sum_{i=1}^{m_1} \left(s_i^{-1} a_i + s_i^{-1} \mathscr{P}_i x\right) = \mathcal{A}^\dagger s^{-1} + \mathcal{P}^\dagger[x] s^{-1},$$

and that (see (1) - (3))

$$\begin{aligned}
\sum_{i=1}^{m_1} \frac{\langle a_i + \mathscr{P}_i x, \xi\rangle \langle a_i + \mathscr{P}_i x, \zeta\rangle}{s_i^2(x)} &= \sum_{i=1}^{m_1} \left(\left\langle s_i^{-2} a_i, \xi\right\rangle \langle a_i, \zeta\rangle + \left\langle s_i^{-2} a_i, \xi\right\rangle \langle \mathscr{P}_i x, \zeta\rangle\right) \\
&\quad + \sum_{i=1}^{m_1} \left(\left\langle s_i^{-2} \mathscr{P}_i x, \xi\right\rangle \langle a_i, \zeta\rangle + \left\langle s_i^{-2} \mathscr{P}_i x, \xi\right\rangle \langle \mathscr{P}_i x, \zeta\rangle\right) \\
&= \left(S^{-2}\mathcal{A}\xi\right)^{\mathsf{T}} \mathcal{A}\zeta + \left(S^{-2}\mathcal{P}[x]\xi\right)^{\mathsf{T}} \mathcal{A}\zeta + \left(S^{-2}\mathcal{A}\xi\right)^{\mathsf{T}} \mathcal{P}[x]\zeta \\
&\quad + \left(S^{-2}\mathcal{P}[x]\xi\right)^{\mathsf{T}} \mathcal{P}[x]\zeta \\
&= \left\langle \mathcal{A}^\dagger S^{-2} \mathcal{A}\xi, \zeta\right\rangle + \left\langle \mathcal{A}^\dagger S^{-2} \mathcal{P}[x]\xi, \zeta\right\rangle + \left\langle \mathcal{P}^\dagger[x] S^{-2} \mathcal{A}\xi, \zeta\right\rangle \\
&\quad + \left\langle \mathcal{P}^\dagger[x] S^{-2} \mathcal{P}[x]\xi, \zeta\right\rangle \\
&= \left\langle \left(\mathcal{A}^\dagger + \mathcal{P}^\dagger[x]\right) S^{-2} \left(\mathcal{A} + \mathcal{P}[x]\right) \xi, \zeta\right\rangle.
\end{aligned}$$

Therefore, from (20) and (21), the first and second Fréchet derivatives of $\eta(x, \mu)$ are, respectively, given by

$$\mathcal{D}_x \eta(x, \mu) = a_0 + \mathscr{P}_0 x + \mu \left(\mathcal{A}^\dagger + \mathcal{P}^\dagger[x]\right) s^{-1} - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} \lambda^{(k)}, \tag{22}$$

and

$$\begin{aligned}
\mathcal{D}_{xx}^2 \eta(x, \mu)(\xi, \zeta) &= \langle \mathscr{P}_0 \xi, \zeta\rangle + \mu \left\langle \mathcal{P}^\dagger[\xi] s^{-1}, \zeta\right\rangle + \mu \left\langle \left(\mathcal{A}^\dagger + \mathcal{P}^\dagger[x]\right) S^{-2} \left(\mathcal{A} + \mathcal{P}[x]\right) \xi, \zeta\right\rangle \\
&\quad - \sum_{k=1}^{K} \left\langle \mathcal{T}^{(k)\dagger} L^{(k)} N^{(k)} L^{(k)} \mathcal{T}^{(k)} \xi, \zeta\right\rangle,
\end{aligned} \tag{23}$$

for any $\xi, \zeta \in \mathbf{H}$, or equivalently

$$\begin{aligned}
\mathcal{D}_{xx}^2 \eta(x, \mu)(\xi) &= \mathscr{P}_0 \xi + \mu \mathcal{P}^\dagger[\xi] s^{-1} + \mu \left(\mathcal{A}^\dagger + \mathcal{P}^\dagger[x]\right) S^{-2} \left(\mathcal{A} + \mathcal{P}[x]\right) \xi \\
&\quad - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} L^{(k)} N^{(k)} L^{(k)} \mathcal{T}^{(k)} \xi,
\end{aligned}$$

for any $\xi \in \mathbf{H}$, where the matrices $L^{(k)}$ and $N^{(k)}, k = 1, 2, \ldots, K$, are defined in (18).

5. Self-concordance analysis

In this section, we show that the recourse function $\eta(\cdot, \cdot)$ is a μ-self-concordant function on $\mathcal{F}_2$. Then we show that the set of all recourse functions forms a strongly self-concordant family with given parameters. These features are of high importance because they yield a nice performance of Newton's method and can guarantee a polynomial time execution of the proposed algorithm. We point out that the corresponding self-concordance analysis of Nesterov and Nemirovskii [16] for the finite-dimensional setting is found in Renegar [17] (see also Renegar [18] and Faybusovich and Moore [11] and Alzalg [4]) for the infinite-dimensional setting.

5.1. Self-concordance of the recourse function

We first introduce the definition of a self-concordant function. Nesterov and Nemirovskii introduced this definition for finite-dimensional optimization (see Definition 2.1.1 in [16]). The definition is introduced below for optimization in a Hilbert space (see also [17]).

Definition 1. Let C be an open nonempty convex subset of a Hilbert space $\mathbf{H}$, and let f be thrice Fréchet differentiable, convex mapping from C to $\mathbb{R}$. Then f is called α-self-concordant on C with a parameter $\alpha > 0$ if for every $x \in C$ and $\xi \in \mathbf{H}$, the following inequality holds

$$\left|\mathscr{D}^3_{xxx} f(x)(\xi,\xi,\xi)\right| \;\leq\; 2\,\alpha^{-1/2}\left(\mathscr{D}^2_{xx} f(x)(\xi,\xi)\right)^{3/2}.$$

An α-self-concordant function f on C is called strongly α-self-concordant if f tends to infinity for any sequence approaching a boundary point of C. The parameter α is called the complexity value of the self-concordant function f.

For any $\mu > 0$, $x \in \mathcal{F}_1$ and $\xi \in \mathbf{H}$, it is easy to verify the following properties of the function $\eta(\cdot,\cdot)$.

Property 1. The function $\eta(x,\mu)$ is continuous on $\mathbb{R}_{++}\times\mathcal{F}_0$ and convex on $\mathcal{F}_0$ for fixed $\mu \in \mathbb{R}_{++}$.

Property 2. The function $\eta(x,\mu)$ has three Fréchet derivatives on $\mathcal{F}_0$, which are continuous on $\mathbb{R}_{++}\times\mathcal{F}_0$ and continuously differentiable in $\mu \in \mathbb{R}_{++}$.

Property 3. Along any sequence $\{x_i \in \mathcal{F}_0\}_{i=1}^{\infty}$ converging to the boundary of $\mathcal{F}_0$, the function $\eta(x,\mu)$ tends to infinity.

Although Properties 1, 2 and 3 of $\eta(x,\mu)$ are essential, they are insufficient to assure that Newton's method used in the proposed algorithms performs well. To prove that the recourse function is a self-concordant map, we first give with a proof the following lemma.

Lemma 1. *For every $\mu > 0$, $x \in \mathcal{F}_2^{(k)}$ and $\xi \in \boldsymbol{H}$, we have that*

$$|\mathcal{D}^3_{xxx}\rho^{(k)}(\xi,\xi,\zeta)| \leq 2\mu^{-1/2}\left(\mathcal{D}^2_{xx}\rho^{(k)}(x,\mu)(\xi,\xi)\right)^{3/2}. \tag{24}$$

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

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

Note that $\phi^{(k)}(0) \triangleq \mathcal{D}^2_{xx}\rho^{(k)}(x,\mu)(\xi,\xi)$ and $\phi^{(k)\prime}(0) \triangleq \mathcal{D}^3_{xxx}\rho^{(k)}(x,\mu)(\xi,\xi,\xi)$. So, to prove that (24) is satisfied for $\rho^{(k)}(x,\mu)$ on $\mathcal{F}_2^{(k)}$, it suffices to show that

$$\left|\phi^{(k)\prime}(0)\right| \leq \frac{1}{\sqrt{\mu}} \left|\phi^{(k)}(0)\right|^{3/2}.$$

Let $(y(t), z^{(k)}(t), \lambda^{(k)}(t)) \triangleq (y(\mu, x + t\xi), z^{(k)}(\mu, x + t\xi), \lambda^{(k)}(\mu, x + t\xi))$, $N^{(k)}(t) \triangleq N^{(k)}(\mu, x + t\xi)$, $L^{(k)}(t) \triangleq L^{(k)}(\mu, x + t\xi)$, and $\mathscr{R}^{(k)}(t) \triangleq \mathscr{R}^{(k)}(\mu, x + t\xi)$. Then

$$(y, z^{(k)}, \lambda^{(k)}) = (y(0), z^{(k)}(0), \lambda^{(k)}(0)),\ L^{(k)} = L^{(k)}(0),\ \mathscr{R}^{(k)} = \mathscr{R}^{(k)}(0),\ \text{and } N^{(k)} = N^{(k)}(0).$$

We define $u^{(k)}(t) \triangleq N^{(k)}(t)L^{(k)}\mathcal{T}^{(k)}(t)\xi$ and $u^{(k)} \triangleq u^{(k)}(0)$.
We can easily show that $N^{(k)^2} = N^{(k)}$. Then using (19) and (17), we have that

$$\begin{aligned}
\phi^{k}(0) &= \mathcal{D}^2_{xx}\rho^{(k)}(\mu,\xi)(\xi,\xi)\\
&= \left\langle \mathcal{T}^{(k)\dagger}\mathcal{J}_x\left[\lambda^{(k)}\right]\xi, \xi\right\rangle\\
&= \sum_{j=1}^{m_2} \left\langle \mathcal{T}^{(k)\dagger}L^{(k)}N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi, \xi\right\rangle\\
&= \sum_{j=1}^{m_2} \left\langle \mathcal{T}^{(k)\dagger}L^{(k)}N^{(k)2}L^{(k)}\mathcal{T}^{(k)}\xi, \xi\right\rangle\\
&= \left(N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi\right)^{\mathsf{T}} N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi \;=\; \left\|u^{(k)}\right\|_2^2,
\end{aligned}$$

hence $\phi^{(k)\prime}(0) = 2u^{(k)^{\mathsf{T}}}u^{(k)\prime}$.
In addition, since $\mathcal{Q}_i, i = 0, 1, \ldots, m_1$, are self-adjoint nonnegative definite bounded operators on $\mathbf{H}$, we have that

$$\begin{aligned}
\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\mathscr{R}^{(k)-1}\xi &= \left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\left(\left(Q^{(k)\dagger}[y] + \mathcal{W}^{(k)\dagger}\right)L^{(k)2}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right) + \mathcal{Q}_0^{(k)} + \sum_{j=1}^{m_2}\lambda_j^{(k)}\mathcal{Q}_j^{(k)}\right)^{-1}\xi\\
&\leq \left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\tilde{\mathscr{R}}^{(k)-1}\xi,
\end{aligned}$$

for any $\xi \in \mathbf{H}$, where $\tilde{\mathscr{R}}^{(k)}\xi \triangleq \left(Q^{(k)\dagger}[y] + \mathcal{W}^{(k)\dagger}\right)L^{(k)2}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi$. Then we have

$$\begin{aligned}
u^{(k)\prime} &= \left\{N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi\right\}'\\
&= \left\{L^{(k)} - L^{(k)}\left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\mathscr{R}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)L^{(k)2}\right\}'\mathcal{T}^{(k)}\xi\\
&\leq \left\{L^{(k)} - L^{(k)}\left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)L^{(k)2}\right\}'\mathcal{T}^{(k)}\xi\\
&= \Big\{L^{(k)\prime} - L^{(k)\prime}\left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)L^{(k)2}\\
&\quad + L^{(k)}\left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)\left(L^{(k)}L^{(k)\prime} + L^{(k)\prime}L^{(k)}\right)\\
&\quad \left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)L^{(k)2}\\
&\quad - \left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)\left(L^{(k)}L^{(k)\prime} + L^{(k)\prime}L^{(k)}\right)\Big\}'\mathcal{T}^{(k)}\xi\\
&= \left\{L^{(k)\prime} - L^{(k)\prime}\left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)\left(L^{(k)}L^{(k)\prime} + L^{(k)\prime}L^{(k)}\right)\right\}\\
&\quad \left(I - \left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\mathscr{R}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)L^{(k)2}\right)\mathcal{T}^{(k)}\xi\\
&= \left\{L^{(k)\prime} - L^{(k)}\left(\mathcal{W}^{(k)} + Q^{(k)}[y]\right)\tilde{\mathscr{R}}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + Q^{(k)\dagger}[y]\right)\left(L^{(k)}L^{(k)\prime} + L^{(k)\prime}L^{(k)}\right)\right\}L^{(k)-1}u^{(k)}.
\end{aligned}$$

Since $L^{(k)}$ is symmetric matrix, it is easy to verify that $N^{(k)}$ is also symmetric. So, $(N^{(k)}v)^{\mathsf{T}}w = v^{\mathsf{T}}N^{(k)}w$ for any $v, w \in \mathbb{R}^{m_2}$, then for any $\xi \in \mathbf{H}$, we have

$$
\begin{array}{cl}
 & u^{(k)^{\mathsf{T}}} L^{(k)} \left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)^{(k)} \xi \\
= & \left(N^{(k)}L^{(k)}\mathcal{T}\xi\right) L^{(k)} \left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi \\
= & \left(L^{(k)}\mathcal{T}\xi\right) N^{(k)}L^{(k)} \left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi \\
= & \left(L^{(k)}\mathcal{T}\xi\right)\left(I - L^{(k)}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\mathscr{R}^{(k)-1}\left(Q^{(k)\dagger}[y] + \mathcal{W}^{(k)\dagger}\right)L^{(k)}\right) \\
 & L^{(k)}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi \\
\leq & \left(L^{(k)}\mathcal{T}\xi\right)\left(I - L^{(k)}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\tilde{\mathscr{R}}^{(k)-1}\left(Q^{(k)\dagger}[y] + \mathcal{W}^{(k)\dagger}\right)L^{(k)}\right) \\
 & L^{(k)}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi \\
= & \left(L^{(k)}\mathcal{T}\xi\right)\left(L^{(k)}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi\right. \\
 & -L^{(k)}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\tilde{\mathscr{R}}^{(k)-1}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right) \\
 & \left. L^{(k)2}\left(Q^{(k)}[y] + \mathcal{W}^{(k)}\right)\xi\right) \\
= & \left(L^{(k)}\mathcal{T}\xi\right)\left(L^{(k)}\mathcal{W}^{(k)}\xi - L^{(k)}\mathcal{W}^{(k)}\mathscr{R}^{(\tilde{k})-1}\tilde{\mathscr{R}}^{(k)}\xi\right) \;=\; 0.
\end{array}
$$

This implies that

$$
\phi^{(k)\prime}(0) \leq \sum_{j=1}^{m_2} 2u^{(k)^{\mathsf{T}}} L^{(k)\prime} L^{(k)-1} u^{(k)}. \tag{25}
$$

By (17), (18), (15) and (25), and using norm inequalities, we have that

$$
\begin{array}{rcl}
|\phi^{(k)\prime}(0)| & \leq & 2\left|u^{(k)^{\mathsf{T}}} L^{(k)\prime} L^{(k)-1} u^{(k)\prime}\right| \\
 & = & \left|u^{(k)^{\mathsf{T}}}\left(L^{(k)\prime}L^{(k)-1} + L^{(k)-1}L^{(k)\prime}\right)u^{(k)}\right| \\
 & = & \left|u^{(k)^{\mathsf{T}}} L^{(k)-1}\left(L^{(k)}L^{(k)\prime} + L^{(k)\prime}L^{(k)}\right)L^{(k)-1}u^{(k)}\right| \\
 & = & \left|u^{(k)^{\mathsf{T}}} L^{(k)-1}\left(L^{(k)^2}\right)' L^{(k)-1}u^{(k)}\right| \\
 & = & \left|u^{(k)^{\mathsf{T}}} \Lambda^{(k)-1}\left(L^{(k)^2}\right)' \Lambda^{(k)-1}u^{(k)}\right| \\
 & = & \left|u^{(k)^{\mathsf{T}}} \Lambda^{(k)-1}\left(\Lambda^{(k)}\Lambda^{(k)\prime} + \Lambda^{(k)\prime}\Lambda^{(k)}\right)\Lambda^{(k)-1}u^{(k)}\right| \\
 & = & 2\left|u^{(k)^{\mathsf{T}}}\left(\Lambda^{(k)-1}\Lambda^{(k)}\Lambda^{(k)\prime}\Lambda^{(k)-1}\right)u^{(k)}\right| \\
 & \leq & 2\left\|u^{(k)}\right\|_2^2\left\|\left(\Lambda^{(k)-1}\Lambda^{(k)\prime}\right)\right\|_2 \\
 & = & 2\mu^{-1/2}\left\|u^{(k)}\right\|_2^2\left\|\left(L^{(k)-1}\mathcal{J}_x\left[\lambda^{(k)}\right]\xi\right)\right\|_2 \text{ (since } L^{(k)-1} = \mu^{1/2}\Lambda^{(k)-1}) \\
 & = & 2\mu^{-1/2}\left\|u^{(k)}\right\|_2^2\left\|\left(L^{(k)-1}L^{(k)}N^{(k)}L^{(k)}\xi\right)\right\|_2 \\
 & = & 2\mu^{-1/2}\left\|u^{(k)}\right\|_2^2\left\|\left(N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi\right)\right\|_2 \left(\text{noting that } u^{(k)} = N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi\right) \\
 & = & 2\mu^{-1/2}\left\|u^{(k)}\right\|_2^2\left\|u^{(k)}\right\|_2 \\
 & = & 2\mu^{-1/2}\left\|u^{(k)}\right\|_2^3 \;=\; 2\mu^{-1/2}\left|\phi^{(k)}(0)\right|^{3/2},
\end{array}
$$

as desired. The proof is now complete.

The following theorem plays a crucial role in this work.

Theorem 1. *For every $\mu > 0$, the recourse function $\eta(x, \mu)$ is a μ-strongly self-concordant function on $\mathcal{F}_0$.*

Proof First, we need to prove that the logarithmic barrier $\ell(x) \triangleq -\mu \sum_{i=1}^{m_1} \ln(-p_i(x))$ is a μ-self-concordant barrier on $\mathcal{F}_1$. Note that the map $\ell(x)$ is convex because $\ell_i(x) \triangleq -\mu \ln(-p_i(x))$, for $i = 1, 2, \ldots, m_1$, is a convex function and the sum of convex functions is convex. Now, due to [11, Section 3], using Taylor expansion of the function $\ell_i(x)$, one can find that the nth Fréchet derivative of $\ell_i(x)$ is

$$\mathcal{D}^n \ell_i(x)(h, h, \ldots, h) = \mu(n-1)! \left(\frac{1}{t_1^n} + \frac{1}{t_2^n}\right),\ n \geq 1,$$

with $t_1^n < 0 < t_2^n$ due to the convexity of $p_i(x + th)$, $t \in \mathbb{R}$. Therefore, the second and third Fréchet derivatives of $\ell_i(x)$ are

$$\begin{aligned} \mathcal{D}^2_{xx}\ell_i(x)(h,h) &= \mu\left(\frac{1}{t_1^2} + \frac{1}{t_2^2}\right), \\ \mathcal{D}^3_{xxx}\ell_i(x)(h,h,h) &= 2\mu\left(\frac{1}{t_1^3} + \frac{1}{t_2^3}\right). \end{aligned}$$

It follows that

$$\begin{aligned} \left|\mathcal{D}^3_{xxx}\ell_i(x)(h,h,h)\right| &= 2\mu\left|\left(\frac{1}{t_1^3} + \frac{1}{t_2^3}\right)\right| \\ &= 2\mu\left|\sum_{i=1}^2 \frac{1}{t_i^3}\right| \\ &\leq 2\mu\left(\sum_{i=1}^2 \frac{1}{t_i^2}\right)^{\frac{3}{2}} \\ &= 2\mu^{-\frac{1}{2}}\left(\mu\sum_{i=1}^2 \frac{1}{t_i^2}\right)^{\frac{3}{2}} \\ &= 2\mu^{-\frac{1}{2}}\left(\left|\mathcal{D}^2_{xx}\ell_i(x)(h,h)\right|\right)^{\frac{3}{2}}, \end{aligned}$$

for $i = 1, 2, \ldots, m_1$. This means that $\ell_i(x)$ is a μ-self-concordant barrier on $\mathcal{F}_1$ for $i = 1, 2, \ldots, m_1$. From [16, Proposition 2.1.1(ii)], we conclude that $\ell(x)$ is a μ-self-concordant barrier on $\mathcal{F}_1$.
In light of Lemma 1, $\rho^{(k)}(\cdot, \cdot)$ is μ-self-concordant on $\mathcal{F}_2^{(k)}$ for $k = 1, 2, \ldots, K$. It can be also seen that $p_0(x)$ is μ-self-concordant on $\mathcal{F}_1$. Using [16, Proposition 2.1.1(ii)] again, we conclude that $\eta(\cdot, \cdot)$ is a μ-self-concordant function on $\mathcal{F}_0$. That is, for every $\mu > 0, x \in \mathcal{F}_0$ and $\xi \in \mathbf{H}$, the following inequality holds.

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

This, together with Property 3, implies that for any fixed $\mu > 0$, the map $\eta(\cdot, \mu)$ is a μ-strongly self-concordant function on $\mathcal{F}_0$. The proof is complete. □

5.2. Parameters of self-concordance family

In this part, we use the result in the preceding subsection to demonstrate that the set of functions $\{\eta(\cdot, \mu) : \mu > 0\}$ is a strongly self-concordant family with appropriate parameters. First, we introduce the concept of a self-concordant family proposed by Renegar [17] for the infinite-dimensional case based on Nesterov and Nemirovskii's definition in [16] for the finite-dimensional case.

Definition 2. Let G be an open nonempty convex subset of Hilbert space $\mathbf{H}$. Let also $\mu \in \mathbb{R}$ and $f_\mu : \mathbb{R}_{++} \times G \to \mathbb{R}$ be a family of functions indexed by μ. Let $\alpha_1(\mu), \alpha_2(\mu), \alpha_3(\mu), \alpha_4(\mu), \alpha_5(\mu) : \mathbb{R}_{++} \to \mathbb{R}_{++}$ be continuously differentiable function on μ. Then the family of functions $f_{\mu \in \mathbb{R}_{++}}$ 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 f_μ is continuous on $\mathbb{R}_{++} \times G$, which are continuous on $\mathbb{R}_{++} \times G$ and continuously differentiable with respect to μ on $\mathbb{R}_{++}$.

(ii) For any $\mu \in \mathbb{R}_{++}$, the function f_μ is strongly $\alpha_1(\mu)$-self-concordant.

(iii) For any $(x, \mu) \in \mathbb{R}_{++} \times G$ and only $h \in \mathbf{H}$

(a) $\left| \frac{\partial}{\partial \mu} \left(\mathcal{D}_x f_\mu(x,\mu)(\xi) \right) - \frac{\partial}{\partial \mu} (\ln \alpha_3(\mu))\, \mathcal{D}_x f_\mu(x,\mu)(\xi) \right| \leq \alpha_4(\mu) \left(\alpha_1(\mu)\right)^{1/2} \left(\mathcal{D}^2_{xx} f_\mu(x,\mu)(\xi,\xi) \right)^{1/2}.$

(b) $\left| D^2_{xx} f_\mu(x,\mu)(\xi,\xi) - \frac{\partial}{\partial \mu} (\ln \alpha_2(\mu))\, \mathcal{D}^2_{xx} f_\mu(x,\mu)(\xi,\xi) \right| \leq \alpha_5(\mu) \mathcal{D}^2_{xx} f_\mu(x,\mu)(\xi,\xi).$

To prove that the set of recourse functions is a self-concordant family, we need first to give with proofs the following two lemmas.

Lemma 2. *For any $\mu > 0, x \in \mathcal{F}_1$ and $\xi \in \boldsymbol{H}$, we have that*

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

Proof By differentiating (13) with respect to μ, we obtain

$$\begin{aligned} \left\langle \mathcal{Q}^{(k)}_j y, y' \right\rangle + \left\langle w^{(k)}_j, y' \right\rangle + z^{(k)\prime}_j &= 0, \quad j = 1, 2, \ldots, m_1, \\ \left\langle y, \mathcal{Q}^{(k)}_0 y' + \sum_{j=1}^{m_2} \lambda^{(k)\prime}_j \left(w^{(k)}_j + \mathcal{Q}^{(k)}_j y \right) + \sum_{j=1}^{m_2} \lambda^{(k)}_j \mathcal{Q}^{(k)}_j y' \right\rangle &= 0, \\ Z^{(k)} \lambda' + \Lambda^{(k)} z' &= e. \end{aligned} \tag{26}$$

Following our notations introduced in Subsection 2, System (26) can be written more compactly as

$$\begin{aligned} \mathcal{Q}^{(k)}[y] y' + \mathcal{W}^{(k)} y' + z' &= 0, \\ \left\langle y,\ \mathcal{Q}^{(k)}_0 y' + \left(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y] \right) \lambda' + \mathcal{Q}^{(k)\dagger}[y'] \lambda^{(k)} \right\rangle &= 0, \\ Z^{(k)} \lambda' + \Lambda^{(k)} z' &= e. \end{aligned}$$

Solving System (26), we get

$$\begin{aligned} y' &= -\mathscr{R}^{(k)-1} \left(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y] \right) z^{(k)-1}, \\ \lambda^{(k)\prime} &= \tfrac{1}{\sqrt{\mu}} L^{(k)} N^{(k)} e, \\ z^{(k)\prime} &= \left(\mathcal{W}^{(k)} + \mathcal{Q}^{(k)}[y] \right) \mathscr{R}^{(k)-1} \left(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y] \right) z^{(k)-1}. \end{aligned} \tag{27}$$

Differentiating (22) with respect to μ and applying (27), we get

$$\{\mathcal{D}_x \eta(x,\mu)\}' = -\frac{1}{\sqrt{\mu}} \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} L^{(k)} N^{(k)} e + \sum_{i=1}^{m_1} \frac{\mathscr{P}_i x + a_i}{s_i} = -\frac{1}{\sqrt{\mu}} \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} L^{(k)} N^{(k)} e + \left(\mathcal{A}^\dagger + \mathcal{P}^\dagger[x] \right) S^{-1} e. \tag{28}$$

Now we define

$$\begin{array}{rclrcl} \mathcal{M}\xi & \triangleq & S^{-1}(\mathcal{A}+\mathcal{P}[x])\,\xi \in \mathbb{R}^{m_1}, & \mathcal{M}^{\dagger}\vartheta & \triangleq & \left(\mathcal{A}^{\dagger}+\mathcal{P}^{\dagger}[x]\right)S^{-1}\vartheta \in \mathbf{H}, \\ \mathcal{M}^{(k)}\xi & \triangleq & -\dfrac{1}{\sqrt{\mu}}N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi \in \mathbb{R}^{m_2}, & \mathcal{M}^{(k)\dagger}v & \triangleq & -\dfrac{1}{\sqrt{\mu}}\mathcal{T}^{(k)\dagger}L^{(k)}N^{(k)}v \in \mathbf{H}, \end{array}$$

and

$$\mathscr{M}\xi \triangleq \sum_{k=1}^{K}\mathcal{M}^{(k)\dagger}\mathcal{M}^{(k)}\xi + \mathcal{M}^{\dagger}\mathcal{M}\xi = \frac{1}{\mu}\sum_{k=1}^{K}\mathcal{T}^{(k)\dagger}L^{(k)}N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi + \left(\mathcal{A}^{\dagger}+\mathcal{P}^{\dagger}[x]\right)S^{-2}(\mathcal{A}+\mathcal{P}[x])\,\xi,$$

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

$$\begin{array}{rcl} \langle \mathscr{M}\xi, \xi\rangle & = & \displaystyle\sum_{k=1}^{K}\left\langle \mathcal{M}^{(k)\dagger}\mathcal{M}^{(k)}\xi, \xi\right\rangle + \mu\left\langle \mathcal{M}^{\dagger}\mathcal{M}\xi, \xi\right\rangle \\ & = & \displaystyle\sum_{k=1}^{K}\left(\mathcal{M}^{(k)}\xi\right)^{\mathsf{T}}\mathcal{M}^{(k)}\xi + \mu\,(\mathcal{M}\xi)^{\mathsf{T}}\mathcal{M}\xi \\ & = & (\Upsilon\xi)^{\mathsf{T}}\,\Upsilon\xi \;=\; \left\langle \Upsilon^{\dagger}\Upsilon\xi, \xi\right\rangle, \end{array}$$

where, for $\xi \in \mathbf{H}, \vartheta \in \mathbb{R}^{m_1}$ and $v_k \in \mathbb{R}^{m_2}, k = 1, 2, \ldots, K$, $\Upsilon\cdot$ and $\Upsilon^{\dagger}\cdot$ are defined as

$$\Upsilon\xi \triangleq \begin{bmatrix} \mathcal{M}^{(1)}\xi \\ \vdots \\ \mathcal{M}^{(K)}\xi \\ \mathcal{M}\xi \end{bmatrix} \in \underbrace{\mathbb{R}^{m_2}\times\cdots\times\mathbb{R}^{m_2}}_{K-\text{times}}\times\mathbb{R}^{m_1}, \text{ and } \Upsilon^{\dagger}\begin{bmatrix} v_1 \\ \vdots \\ v_K \\ \vartheta \end{bmatrix} \triangleq \sum_{k=1}^{K}\mathcal{M}^{(k)\dagger}v_k + \mathcal{M}^{\dagger}\vartheta \in \mathbf{H}.$$

Observe that the operator $\mathscr{M}\cdot = \Upsilon^{\dagger}\Upsilon\cdot$ is invertible from the Hilbert space $\mathbf{H}$ into itself, and hence its inverse operator $\mathscr{M}^{-1}\cdot = \left(\Upsilon^{\dagger}\Upsilon\right)^{-1}\cdot$ is well-defined on $\mathbf{H}$.
Because $\mathscr{P}_i, i = 0, 1, \ldots, m_1$, are self-adjoint nonnegative definite bounded operators on $\mathbf{H}$, we also observe that

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

and that

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

where

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

Using (29) and (30), we obtain

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

By (31) and using norm inequalities, we get

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

The proof is complete. □

Lemma 3. *For any $\mu \ge 0, x \in \mathcal{F}_1$ and $\xi \in \mathbf{H}$, we have that*

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

Proof Let $(z^{(k)},\lambda^{(k)},L^{(k)},\mathscr{R}^{(k)},N^{(k)}) \triangleq (z^{(k)}(x,\mu),\lambda^{(k)}(x,\mu),L^{(k)}(x,\mu),\mathscr{R}^{(k)}(x,\mu),N^{(k)}(x,\mu))$, fix $\xi \in \mathbf{H}$, and define $u^{(k)} \triangleq N^{(k)}L^{(k)}\mathcal{T}^{(k)}\xi$. Then we have that

$$\mathcal{D}^2_{xx}\eta(x,\mu)(\xi,\zeta) = \langle\mathscr{P}_0\xi,\zeta\rangle + \mu\left\langle\mathcal{P}^\dagger[\xi]s^{-1},\zeta\right\rangle + \mu\left\langle\left(\mathcal{A}^\dagger+\mathcal{P}^\dagger[x]\right)S^{-2}\left(\mathcal{A}+\mathcal{P}[x]\right)\xi,\zeta\right\rangle - \sum_{k=1}^{K}u^{(k)^\mathsf{T}}u^{(k)}.$$

Using Lemma 1, we also have

$$\left\{\mathcal{D}^2_{xx}\eta(x,\mu)(\xi,\xi)\right\}' = \left\langle\mathcal{P}^\dagger[\xi]s^{-1},\zeta\right\rangle + \left\langle\left(\mathcal{A}^\dagger+\mathcal{P}^\dagger[x]\right)S^{-2}\left(\mathcal{A}+\mathcal{P}[x]\right)\xi,\zeta\right\rangle - \sum_{k=1}^{K}u^{(k)^\mathsf{T}}L^{(k)}\left(L^{(k)-2}\right)'L^{(k)}u^{(k)}.$$

Note that

$$\begin{aligned} u^{(k)^\mathsf{T}}L^{(k)}\left(L^{(k)-2}\right)'L^{(k)}u^{(k)} &= u^{(k)^\mathsf{T}}\left(Z^{(k)-1}\Lambda^{(k)}\right)^{-1/2}\left(Z^{(k)-1}\Lambda^{(k)}\right)'\left(Z^{(k)-1}\Lambda^{(k)}\right)^{1/2}u^{(k)} \\ &= u^{(k)^\mathsf{T}}\left(\mu^{-1}Z^{(k)2}\right)^{-1/2}\left(Z^{(k)-1}\Lambda^{(k)}\right)'\left(\mu^{-1}Z^{(k)2}\right)^{-1/2}u^{(k)} \\ &= u^{(k)^\mathsf{T}}\left(Z^{(k)2}\right)^{-1/2}\left(\Lambda^{(k)-1}Z^{(k)\prime} - Z^{(k)}\Lambda^{(k)-2}\Lambda^{(k)\prime}\right)\left(z^{(k)2}\right)^{-1/2}u^{(k)} \\ &= \mu^{-1}u^{(k)^\mathsf{T}}\left(\Lambda^{(k)}Z^{(k)\prime} - Z^{(k)}\Lambda^{(k)\prime}\right)u^{(k)} \\ &= \mu^{-1}u^{(k)^\mathsf{T}}\left(2\Lambda^{(k)}Z^{(k)\prime} - I\right)u^{(k)} \quad \left(\text{using } Z^{(k)}\lambda^{(k)\prime} + \Lambda^{(k)}z^{(k)\prime} = e \text{ from (26)}\right) \\ &= \mu^{-1}u^{(k)^\mathsf{T}}\left(2\mu Z^{(k)-1}Z^{(k)\prime} - I\right)u^{(k)} \\ &\le \mu^{-1}\left\|u^{(k)}\right\|_2^2\left\|I - 2\mu z^{(k)-1}z^{(k)\prime}\right\|_2 \quad \left(\text{using definition of } z^{(k)\prime} \text{ from (27)}\right) \\ &= \left\|u^{(k)}\right\|_2^2\left\|e - 2\mu Z^{(k)-1}\left(\mathcal{W}^{(k)} + \mathcal{Q}^{(k)}[y]\right)\mathscr{R}^{(k)-1}\left(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y]\right)z^{-1}\right\|_2 \\ &= \mu^{-1}\left\|u^{(k)}\right\|_2^2\left\|1 - 2N^{(k)}e\right\|_2 \le \frac{\sqrt{m_2}}{\mu}\left\|u^{(k)}\right\|_2^2, \end{aligned}$$

where the last inequality follows from (26) and from the fact that $\|I - 2N^{(k)}\|_2 \leq 1$, which is due to $I - 2N^{(k)} \preceq I$.
Differentiating $\eta(x,\mu)(\xi,\xi)$ with respect to μ, we get

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

The proof is complete. □

The result in the following theorem is now immediate and it is the most important result in this work.

Theorem 2. *The family of function $\{\eta(\cdot,\mu) : \mu > 0\}$ is a strongly self-concordant family with the following parameters:*

$$\alpha_1(\mu) = \mu,\ \alpha_2(\mu) = \alpha_3(\mu) = 1,\ \alpha_4(\mu) = \frac{\sqrt{m_1 + km_2}}{\mu},\ \alpha_5(\mu) = \frac{\sqrt{m_2}}{2\mu}.$$

Proof Condition (i) in Definition 2 is satisfied by Property 3. Theorem 1 satisfies Condition (ii) and Lemmas 2 and 3 prove Inequalities (iiia) and (iiia). The result is established. □

Families of functions having a property similar to that in Theorem 2 have nice features, Most importantly, the parameters defining the self-concordant family allow us to relate the rate at which the barrier parameter μ is varied and the number of Newton steps required to maintain the proximity to the central path of the perturbed problem. The results in this section means that we can generalize the logarithmic-barrier decomposition interior-point algorithm for the finite-dimensional case to the infinite-dimensional case and keep a polynomial time execution of the proposed algorithm. We will see that the obtained complexity estimates are similar to those in the finite-dimensional setting.

6. The proposed algorithm

In Section 5, we have established that the recourse functions $\eta(\cdot,\mu)$ form a strongly self-concordant family. We have also calcualted the Fréchet derivatives $\mathcal{D}_x\eta(x,\mu)$ and $\mathcal{D}^2_{xx}\eta(x,\mu)$ in Section 4, which are required for computing the Newton direction:

$$\begin{aligned}
\triangle x &\triangleq -\{\nabla^2_{xx}\eta(x,\mu)\}^{-1}\left(\mathcal{D}_x\eta(x,\mu) - \xi^*(x)\right)\\
&= -\{\nabla^2_{xx}\eta(x,\mu)\}^{-1}\left(\left(a_0 + \mathscr{P}_0 x + \mu\sum_{i=1}^{m_1}\frac{a_i + \mathscr{P}_i x}{s_i(x,\mu)}\right) - \xi^*(x)\right),
\end{aligned} \tag{32}$$

where $\xi^*(x)$ is the unique vector in $\mathbf{G}^\perp$ such that $\triangle x \in \mathbf{G}$.
We can determine the measure of proximity of the current point x to the central path by

$$\delta(x,\mu) \triangleq \sqrt{-\mu^{-1}\mathcal{D}^2_{xx}\eta(x,\mu)(\triangle x, \triangle x)}. \tag{33}$$

Note that ξ^* is the solution to the minimization problem

$$\begin{array}{ll} \min & \left\{\nabla^2_{xx}\eta(x,\mu)\right\}^{-1}(\mathcal{D}_x\eta(x,\mu)-\xi, \mathcal{D}_x\eta(x,\mu)-\xi) \\ \text{s.t.} & \xi \in \mathbf{G}^\perp. \end{array} \tag{34}$$

The solution of the minimization problem (34) is unique and can be characterized by

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

hence $\xi^{\mathsf{T}}(x) = \bar{\xi} \in \mathbf{G}^\perp$.
Note also that $\delta(\cdot,\cdot)$ in (28) vanishes at (x,μ) if and only if

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

provided that $(x, s; y^{(1)}, z^{(1)}; \ldots; y^{(K)}, z^{(k)})$ is a feasible solution for (14). Based on the self-concordance analysis established in the preceding section, we propose a logarithmic-barrier path-following interior-point decomposition algorithm for the infinite-dimensional two-stage SQCCQP problem; see Algorithm 1 and Figure 1.
We initialize Algorithm 1 with $x_0 \in \mathcal{F}_1$ as an initial first-stage feasible solution and $\mu_0 > 0$ as an initial value for the barrier parameter, and start it with $\epsilon > 0$ as the required accuracy of the final solution, and γ as the reduction parameter. We also use β as a threshold for calculating the distance (which is δ) between the current point x and the central path. If the current x is too far from the central path, i.e., $\delta > \beta$, we use Newton's method to identify a location near to the central path. The value of μ is then lowered by a factor γ, and the procedure is repeated until the value of μ falls within the tolerance ϵ. We can trace the central path as μ approaches to zero in order to find a strictly feasible ϵ-optimal solution to Problem (10).

7. Complexity analysis

This part is devoted to presenting with proofs the time complexity of Algorithm 1. Note that the proposed algorithm can be branched into two versions based on the selection of γ: The short-step algorithm and long-step algorithm. In the short-step algorithm, the barrier parameter μ is decreased by a factor $\gamma \triangleq 1 - \delta/(m_1 + km_2)^{1/2}$, with $\delta < 0.1$ in each iteration. For the long-step algorithm, the barrier parameter is

Algorithm 1: A logarithmic-barrier path-following interior-point decomposition algorithm for the infinite-dimensional two-stage SQCCQP problem (10) and (11)

input $\epsilon,\ \theta,\ \gamma,\ \mu^0, \beta,\ x^0$
ensure $\epsilon > 0,\ \theta > 0,\ \gamma \in (0, 1),\ \mu^0 > 0,\ \beta > 0,\ x^0 \in \mathcal{F}_0$
initialize $\mu = \mu^0,\ x = x^0$
while $\mu \geq \epsilon$ **do**
    for $k = 1, 2, \ldots, K$ **do**
        solve (13) to obtain $y, z^{(k)}, \lambda^{(k)}$
    compute $\mathcal{D}_x\eta(x, \mu)$ and $\mathcal{D}^2_{xx}\eta(x, \mu)$ using (22) and (23)
    compute $\xi^*(x)$ by solving (34)
    compute $\triangle x$ using (32)
    compute $\delta(x, \mu)$ using (33)
    while $\delta > \beta$ **do**
        set $x = x + \theta \triangle x$
        for k=1,2,. . . ,K **do**
            solve (13) to obtain $(y, z^{(k)}, \lambda^{(k)})$
        compute $\mathcal{D}_x\eta(x, \mu)$ and $\mathcal{D}^2_{xx}\eta(x, \mu)$ using (22) and (23)
        compute $\xi^*(x)$ by solving (34)
        compute $\triangle x$ using (32)
        compute $\delta(x, \mu)$ using (33)
    set $\mu = \gamma\mu$

decreased by the constant factor $\gamma \in (0, 1)$ which is independent of m_1, m_2 and K. We want to determine an upper bound on Itr, which is the number of Newton iterations needed to find the point x^k, such that at each complete iteration, the algorithm performs an outer iteration which updates the parameter μ by the factor γ, this is followed by an inner loop involving several inner Newton iterations. Hence, the total number of Newton iterations needed by the algorithm is not more than

$$\text{Itr} = \text{Itr}_{\text{out}} \times \text{Itr}_{\text{in}},$$

where Itr_{out} and Itr_{in} are upper bounds on the number of iterations performed by the outer while loop (which reduce the parameter μ) and the number of iterations performed by the inner while loop, respectively.
First, we want to estimate Itr_{out} for both short- and long-step algorithms. Let μ^k be the parameter at kth outer iteration, then we have

$$\mu^k = \gamma\mu^{k-1} = \cdots = \gamma^k \mu^0,$$

where γ is the update factor. Then $\mu^k < \epsilon$ if

$$\gamma^k \mu^0 \leq \epsilon,$$

or equivalently

$$k \geq \left(\ln\left(\frac{\mu^0}{\epsilon}\right)\right) \Big/ \ln \gamma^{-1},$$

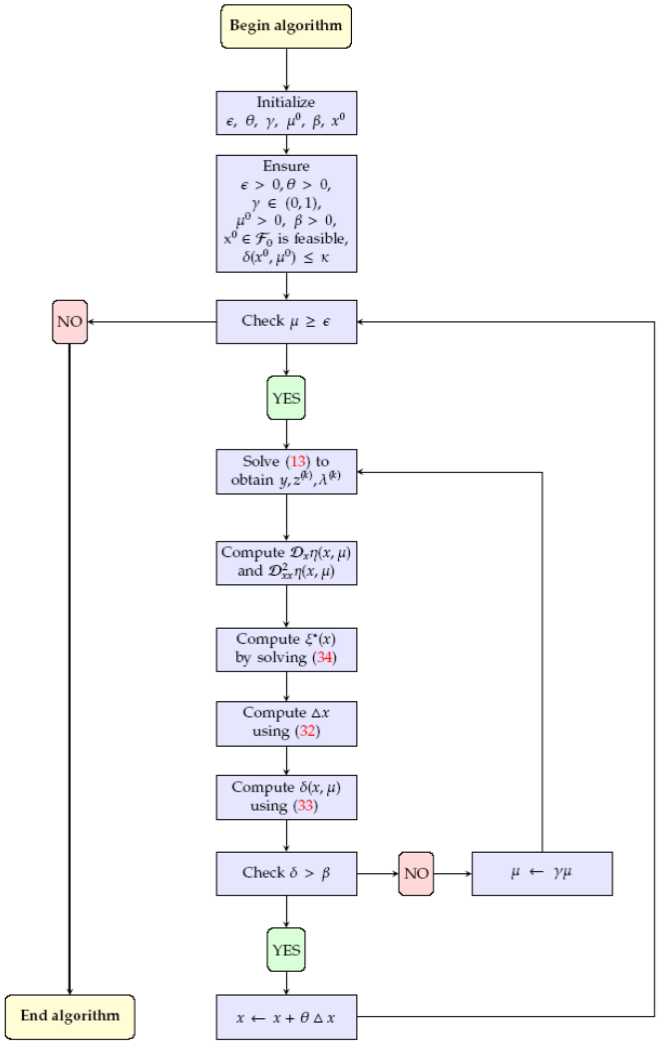

Figure 1. **A flowchart of Algorithm 1.**

and hence

$$\mathrm{Itr}_{\mathrm{out}} \leq \left(\ln\left(\frac{\mu^0}{\epsilon}\right)\right) \Big/ \ln \gamma^{-1} + 1.$$

For short-step algorithm, since $\gamma = \ln(1 - \sigma/\sqrt{m_1 + km_2}) \approx -\sigma/\sqrt{m_1 + km_2}$, we have

$$\mathrm{Itr}_{\mathrm{out}} = \mathcal{O}(1)\sqrt{m_1 + Km_2}\ln\frac{\mu^0}{\epsilon}.$$

For long-step algorithm, since $\gamma = \mathcal{O}(1)$ is a constant, we have

$$\mathrm{Itr}_{\mathrm{out}} = \mathcal{O}(1) \ln \frac{\mu^0}{\epsilon}.$$

Now, we need to estimate the number of inner iterations of Newton's method using two different merit function to measure the agreement of the progress of Newton's iterates. If the agreement is good, the merit function is small. We use $\delta(x, \mu)$ for the short-step algorithm and the first-stage objective function $\eta(x, \mu)$ for the long step algorithm. The following lemma estimates the reduction of the merit function, which corresponds to [16, Theorem 2.2.3] in the finite-dimensional case. This lemma is based on [11] for the infinite-dimensional case. More precisely, Item (i) is due to [11, Proposition 3.4] and Item (ii) is due to [11, Proposition 3.3].

Lemma 4. *For any $\mu > 0$ and $x \in \mathcal{F}_1$, let $x^+ \triangleq x + \triangle x$, $\triangle x^+$ be the Newton direction calculated at x^+ and $\delta(\mu, x^+) \triangleq \sqrt{-\frac{1}{\mu}\mathcal{D}^2_{xx}\eta(x,\mu)\,(\triangle x^+, \triangle x^+)}$. Then the following statements hold:*

(i) If $\delta < 2 - \sqrt{3}$, then $\delta(\mu, x^+) \leq \delta/2$.

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

In the remaining part of this section, we estimate the upper bound on number of inner iterations of Newton's method.

Complexity of short-step algorithm The short-step algorithm is executed as follows. At the beginning of the kth iteration, x^k is close to the central path because it satisfies $\delta(\mu^k, x) \leq \beta$. After reducing the parameter μ to $\mu^{(k+1)} = \gamma\mu$, we have that $\delta(\mu^k, x) \leq 2\beta$. Then one Newton step with step size $\theta = 1$ is taken to construct a new point x^{k+1} with $\delta(\mu^k, x) \leq \beta$. We present the complexity result of the short-step algorithm in the following theorem, for which the proof is based on Lemma 5 given in Appendix A.

Theorem 3. *Consider Algorithm 1. Let μ^0 be an initial barrier parameter, $\epsilon > 0$ be the stopping criterion, and $\beta = (2 - \sqrt{3})/2$. If the starting point x^0 is sufficiently close to the central path, i.e., $\delta(\mu^0, x^0) \leq \beta$, then the short-step reduces the parameter μ at linear rate and terminates within*

$$\mathcal{O}\left(\sqrt{m_1 + Km_2} \ln\left(\frac{\mu^0}{\epsilon}\right)\right).$$

Proof Using Item (i) in Lemma 4, and Lemma 5, we find that Algorithm 1 reduces the parameter μ by the factor $\gamma = 1 - 0.1/\sqrt{m_1 + km_2}$ at each iteration, and requires only one Newton step in each inner loop (i.e., $\mathrm{Itr}_{\mathrm{in}} = 1$). The proof complete.

Complexity of long-step algorithm Because iterates generated by the long-step algorithm may be distant from $x(\mu)$, the condition $\delta < 2 - \sqrt{3}$ could be violated, which means that only Item (ii) in Lemma 4 can be used, hence we use η as our merit function. The long-step algorithm is executed as follows. At the beginning of the kth iteration, we have a point x^{k-1} that is sufficiently near to $x(\mu^{k-1})$ (which is the solution to (10) for $\mu \triangleq \mu^{k-1}$). When the barrier parameter is reduced from μ^{k-1} to $\mu^k \triangleq \gamma\mu^{k-1}$, where $\gamma \in (0, 1)$, we search for a point x^k that is sufficiently close to $x(\mu^k)$. This produces a finite sequence of points $p_1, p_2, \dots, p_{\mathrm{Itr}} \in \mathcal{F}_1$. We take x^k to be p_{Itr}. So we need to determine an upper bound on Itr, the total number of Newton iterations. Let

$$\phi \triangleq \eta(x, \mu) - \eta(\mu, x(\mu))$$

be the difference between the minimum objective value $\eta(\mu^k, x(\mu^{k-1}))$ at the beginning of the kth iteration and the objective value $\eta(\mu^k, x(\mu^k))$ at the end of kth iteration. We present the complexity result of the long-step algorithm in the following theorem.

Theorem 4. *Consider Algorithm 1. Let μ^0 be an initial parameter, $\epsilon > 0$ be the stopping criterion, and $\beta = 1/6$. If the starting point x^0 is sufficiently close to the central path, i.e., $\delta(\mu^0, x^0) \le \beta$, then the long-step algorithm reduces the parameter μ at linear rate and terminates within*

$$\mathcal{O}\left(m_1 + Km_2 \ln\left(\frac{\mu^0}{\epsilon}\right)\right).$$

Proof First we need to find an upper bound on $\phi(\mu^+, x)$. Let $\mu^+ \triangleq \gamma\mu$ with $\gamma \in (0, 1)$ and define

$$\tilde{\delta} \triangleq \tilde{\delta}(x, \mu) \triangleq \sqrt{-\frac{1}{\mu}\mathcal{D}^2_{xx}\eta(x, \mu)(\tilde{\Delta}x^+, \tilde{\Delta}x^+)}.$$

We show that if $\tilde{\delta} < 1$, then

$$\eta(\mu^+, x(\mu^+)) - \eta(\mu^+, x) \le \mathcal{O}(m_1 + Km_2)\mu^+. \tag{35}$$

Note that

$$\phi(\mu, x) \triangleq \eta(\mu, x(\mu)) - \eta(x, \mu) = \int_0^1 \mathcal{D}_x\eta(\mu, x + \tau\tilde{\Delta}x)(\tilde{\Delta}x^+)d\tau.$$

Since $x(\mu)$ is the optimal solution, we have

$$\mathcal{D}_x\eta(\mu, x(\mu)) = 0. \tag{36}$$

Then, for any $\mu > 0$, by applying chain rule, using (36) and applying Mean-Value Theorem, we get

$$\begin{aligned} \phi'(x, \mu) &= \eta'(\mu, x(\mu)) - \eta'(x, \mu) - \mathcal{D}_x\eta(\mu, x(\mu))(x'(\mu)) \\ &= \eta'(\mu, x(\mu)) - \eta'(x, \mu) \\ &= \mathcal{D}_x\eta(\mu, x(\mu) + \omega\tilde{\Delta}x)(\tilde{\Delta}x). \end{aligned} \tag{37}$$

Now, we differentiate (37) with respect to μ to get

$$\phi''(x,\mu) = \eta''(x,\mu) - \eta''(\mu, x(\mu)) - \mathcal{D}_x\eta(\mu, x(\mu))(x'(\mu)). \tag{38}$$

We want to bound the terms $-\mathcal{D}_x\eta(\mu, x(\mu)(x'(\mu))$ and $\phi''(x,\mu)$ in (38). By differentiating $\rho^{(k)}(x,\mu)$ with respect to μ and using (28) and (15), we get

$$\begin{aligned}
\rho^{(k)\prime}(x,\mu) &= \langle \mathcal{Q}_0^{(k)} y, y'\rangle + \langle w_0^{(k)}, y'\rangle - e^{\mathsf{T}} \ln z^{(k)} - \mu e^{\mathsf{T}} Z^{(k)-1} z^{(k)\prime} \\
&= \langle \mathcal{Q}_0^{(k)} y + w_0^{(k)}, y'\rangle - e^{\mathsf{T}} \ln z^{(k)} - \Lambda^{(k)} z^{(k)\prime} \ \ \text{(using (13))} \\
&= \langle \mathcal{Q}_0^{(k)} y + w_0^{(k)}, y'\rangle - e^{\mathsf{T}} \ln z^{(k)} \\
&\quad -\Lambda^{(k)}\Big(\mathcal{W}^{(k)} + \mathcal{Q}^{(k)}[y]\Big)\mathcal{R}^{(k)-1}\Big(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y]\Big) z^{(k)-1} \ \ \text{(using (27))} \\
&= \Big\langle \mathcal{Q}_0^{(k)} y + w_0^{(k)}, y'\Big\rangle - e^{\mathsf{T}} \ln z^{(k)} \\
&\quad +\Lambda^{(k)}\Big(\mathcal{W}^{(k)} + \mathcal{Q}^{(k)}[y]\Big)\Big(-\mathcal{R}^{(k)-1}\Big(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y]\Big) z^{(k)-1}\Big) \\
&= \Big\langle \mathcal{Q}_0^{(k)} y + w_0^{(k)}, y'\Big\rangle - e^{\mathsf{T}} \ln z^{(k)} + \Lambda^{(k)}\Big(\mathcal{W}^{(k)} + \mathcal{Q}^{(k)}[y]\Big) y' \\
&= \Big\langle \mathcal{Q}_0^{(k)} y + w_0^{(k)}, y'\Big\rangle - e^{\mathsf{T}} \ln z^{(k)} + \left\langle \sum_{j=1}^{m_2} \Big(\lambda^{(k)} w_j^{(k)} + \lambda^{(k)} \mathcal{Q}_j^{(k)} y\Big), y' \right\rangle \\
&= -\sum_{j=1}^{m_2} \ln z_j^{(k)}.
\end{aligned}$$

Note that $N^{(k)2} = N^{(k)}$. Now, differentiating $\rho^{(k)\prime}$ with respect to μ and using (28) and (15), we get

$$\begin{aligned}
\rho^{(k)\prime\prime}(\mu, x(\mu)) &= e^{\mathsf{T}} Z^{(k)-1} z^{(k)\prime} \\
&= e^{\mathsf{T}} Z^{(k)-1}\Big(\mathcal{W}^{(k)} + \mathcal{Q}^{(k)}[y]\Big)\mathcal{R}^{(k)-1}\Big(\mathcal{W}^{(k)\dagger} + \mathcal{Q}^{(k)\dagger}[y]\Big) z^{(k)-1} \\
&= e^{\mathsf{T}} Z^{(k)-1} L^{(k)-1} N^{(k)} L^{(k)-1} z^{(k)-1}.
\end{aligned}$$

Thus we have

$$\rho^{(k)\prime\prime}(\mu, x(\mu)) \le e^{\mathsf{T}} z^{(k)-1} L^{(k)-2} z^{(k)-1} = e^{\mathsf{T}} z^{(k)-1} \lambda^{(k)-1} = \frac{m_2}{\mu},$$

and hence

$$\rho^{(k)\prime\prime}(\mu, x(\mu)) \le \frac{K m_2}{\mu}. \tag{39}$$

We differentiate the optimality condition of the first-stage problem and get

$$x'(\mu) = -\left\{\nabla_{xx}^2 \eta(\mu, x(\mu))\right\}^{-1} \Big(\{\mathcal{D}_x\eta(\mu, x(\mu))\}'\Big). \tag{40}$$

It follows that

$$\begin{aligned}
\{\mathcal{D}_x\eta((\mu, x(\mu))\}' (x'(\mu)) &= -\left\{\nabla_{xx}^2 \eta(x(\mu), \mu)\right\}^{-1} (\{\mathcal{D}_x\eta(\mu, x(\mu))\}', \{\mathcal{D}_x\eta(\mu, x(\mu))\}') \\
&\le \mu^{-1}(m_1 + K m_2).
\end{aligned} \tag{41}$$

Combining (39) and (41) and using $\eta''(\mu, x) \geq 0$ we have

$$\phi''(x, \mu) \leq (\mu^{-1}(m_1 + 2Km_2)).$$

Now, using (41), and applying Mean-Value Theorem, we have

$$\begin{aligned}\phi(\mu^+, x) &= \phi(x, \mu) + \phi'(x, \mu)(\mu^+ - \mu) + \int_{\mu}^{\mu^+} \int_{\mu}^{\tau} \phi''(v, x) dv d\tau \\ &\leq \mu\left(\tfrac{\tilde{\delta}}{1-\tilde{\delta}} + \ln(1-\tilde{\delta})\right) - \sqrt{m_1 + Km_2}\ln(1-\tilde{\delta})(\mu^+ - \mu) \\ &\quad + (m_1 + 2Km_2)(\mu^+ - \mu)\ln\gamma^{-1}.\end{aligned}$$

Table 1. **The short-step algorithm versus the long-step algorithm.**

Features	Short-step algorithm	Long-step algorithm
Factor γ	$\gamma = 1 - \delta/\sqrt{m_1 + m_2},\ \delta > 0$	$\gamma \in (0, 1)$
Merit function	$\delta(x, \mu)$	$\eta(x, \mu)$
Number of inner iterations	$\mathrm{Itr_{in}} = 1$	$\mathrm{Itr_{in}} = O(1)(m_1 + Km_2)$
Number of outer iterations	$\mathrm{Itr_{out}} = O(1)\sqrt{m_1 + km_2}\ln(\mu^0/\epsilon)$	$\mathrm{Itr_{out}} = \ln(\mu^0/\epsilon)$

As we have already shown, Algorithm 1 reduces the barrier parameter, μ, $\mathrm{Itr_{out}}$ times, such that $\mathrm{Itr_{out}} = O(\ln(\mu^0/\epsilon))$, each outer iteration has several inner Newton's iterations. Now we are looking for an upper bound of $\mathrm{Itr_{in}}$. Assume that $\eta(\mu^{k-1}, x^{k-1}) \leq \beta$ and $\mu^k = \gamma\mu^{k-1}$, and let

$$\tilde{x}^0 \triangleq x^{k-1}, \tilde{x}^1 \triangleq x^k, \ldots, \tilde{x}^j, \ldots$$

be the inner iterates that are generated in the kth loop. We choose β to be small enough so that $\tilde{\delta} \leq \tilde{\beta} < 1$ at (μ^{k-1}, x^{k-1}) (see Lemma 8). From (35), there is a positive constant v such that

$$\phi(\mu^k, x^{k-1}) = \eta(\mu^k, x^{k-1}) - \eta(\mu^k, x(\mu^k)) \leq v(m_1 + Km_2)\mu^k.$$

Assume that $\delta^i \triangleq \delta(\mu^k, \tilde{x}^i) > \beta$, for all $i = 1, \ldots, j-1$, and denote $\delta \triangleq \beta - \ln(1+\beta) > 0$. Then

$$\delta < \delta^i - \ln(1 + \delta^i),\ i = 0, 1, \ldots, j-1.$$

Note that, from Item (ii) in Lemma 4, we have

$$\eta(\mu^k, \tilde{x}^{i+1}) \leq \eta(\mu^k, \tilde{x}^i + \theta \vartriangle x) \leq \eta(\mu^k, \tilde{x}^i) - (\delta^i - \ln(1+\delta^i))\mu^k \leq \eta(\mu^k, \tilde{x}^i) - \sigma\mu^k.$$

As a result

$$\eta(\mu^k, x(\mu)) \leq \eta(\mu^k, \tilde{x}^i) \leq \eta(\mu^k, x^{k-1}) - j\sigma\mu^k + \eta(\mu^k, x(\mu)).$$

Therefore $j \leq \epsilon(m_1 + km_2)/\sigma$, which means that $\sigma \leq \beta$ for any $j > (m_1 + km_2)\sigma$ iterations. Since σ is a constant, we have $\mathrm{Itr_{in}} = \mathscr{O}(m_1 + km_2)$. The proof complete.

Theorems 3 and 4 are the counterparts of Theorems 4.1 and 4.2 in [15] as well as Theorems 6.1 and 6.2 in [4]. One can see that the complexity results in Theorems 3 and 4 coincide with the best known ones for the finite-dimensional case. This confirms that the time execution of Algorithm 1 is independent of the underlying Hilbert space. We end this section with Table 1, which compares between the short- and long-step versions of Algorithm 1.

8. An illustrative example

In this section, we give an example from stochastic control theory and solve it using the algorithm proposed in Section 6. The notations in this section might be a bit different from those in the other sections. The stochastic model under interest extends the control design model problem (see [11], for example) by applying a relaxation on its assumptions that include deterministic data to include random data.
We let $\mathbf{H} \triangleq L_2^m[0,T]$ be set of all measurable square integrable functions on $[0,T]$ with values in $\mathbb{R}^m$. For $i = 1, 2, \ldots, m_1$, we define

$$p_i(x,u) \triangleq \frac{1}{2}\int_0^T \left(x^{\mathsf{T}}(t)\mathscr{P}_i(t)x(t) + u^{\mathsf{T}}(t)\mathscr{R}_i(t)u(t)\right) dt + b_i,$$

where $\mathscr{P}_i(t)$ (respectively, $\mathscr{R}_i(t)$) is an $m_1 \times m_1$ (respectively, $p_1 \times p_1$) symmetric positive definite matrix which depends continuously on $t \in [0,T]$, for $i = 0, 1, \ldots, m_1$, and $b \in \mathbb{R}^{m_1}$. The control objective function is to find (x,u) that minimizes the cost criterion function given by

$$p_0(x,u) \triangleq \frac{1}{2}\int_0^T \left(x^{\mathsf{T}}(t)\mathscr{P}_0(t)x(t) + u^{\mathsf{T}}(t)\mathscr{R}_0(t)u(t)\right) dt.$$

For $j = 0, 1, \ldots, m_2$, we also define

$$q_j(y,v,\omega) \triangleq \frac{1}{2}\int_0^T \left(y(t)^{\mathsf{T}}\mathscr{Q}_j^{(k)}(t,\omega)y(t) + v(t)^{\mathsf{T}}\mathscr{S}_j^{(k)}(t,\omega)v(t)\right) dt,$$

and $\mathscr{Q}_j(t,\omega)$) (respectively, $\mathscr{S}_j(t,\omega)$)) is an $m_2 \times m_2$ (respectively, $p_2 \times p_2$) symmetric positive definite random matrix which depends continuously on $t \in [0,T]$ and its randomness depends on an underling outcome ω in an event space Ω with a known probability function $\mathbb{P}$, for $j = 0, 1, \ldots, m_2$.

We consider the following closed subspaces of $\mathbf{H}$:

$$\mathbf{G} \triangleq \left\{ (x,u) \in L_2^{m_1}[0,T] \times L_2^{p_1}[0,T] : \begin{array}{l} x \text{ is absolutely continuous on } [0,T], \\ \dot{x} \in L_2^{m_1}[0,T], x(0) = 0, \\ \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ \text{for all } t \in [0,T] \end{array} \right\},$$

$$\mathbf{G}(\omega) \triangleq \left\{ (y,v) \in L_2^{m_2}[0,T] \times L_2^{p_2}[0,T] : \begin{array}{l} y \text{ is absolutely continuous on } [0,T], \\ \dot{y} \in L_2^{m_2}[0,T], y(0) = 0, \\ \dot{y}(t) = W(t,\omega)y(t) + D(t,\omega)v(t) \\ \text{for all } t \in [0,T] \end{array} \right\},$$

where $\dot{x}(t)$ and $\dot{y}(t)$ are the time derivatives of the functions $x(\cdot)$ and $y(\cdot)$ with respect to time t, $A(t)$ (respectively, $B(t)$) is an $m_1 \times m_1$ (respectively, $m_1 \times p_1$) matrix which depends continuously on $t \in [0,T]$, and $W(t,\omega)$ (respectively, $D(t,\omega)$) is an $m_2 \times m_2$ (respectively, $m_2 \times p_2$) random matrix which depends continuously on $t \in [0,T]$ and its randomness depends on an underling outcome ω in an event space Ω with a known probability function $\mathbb{P}$.
We are interested in a problem of the form:

$$\begin{array}{ll} \min & p_0(x,u) + \mathbb{E}[\rho(y,v,\omega)] \\ \text{s.t.} & p_i(x,u) \leq 0, \; i = 1,2,\ldots,m_1, \\ & (x,u) \in \mathbf{G}, \end{array} \tag{42}$$

where $\rho(\omega,x,u)$ is the minimum value of a problem of the form:

$$\begin{array}{ll} \min & q_0(y,v,\omega) \\ \text{s.t.} & q_j(y,v,\omega) + \displaystyle\int_0^T \left(w_j^{\mathsf{T}}(\omega)y(t) + w_j(\omega)^{\mathsf{T}}v(t)\right)dt \leq h_j(\omega) - \int_0^T \left(t_j^{\mathsf{T}}(\omega)x(t) + t_j^{\mathsf{T}}(\omega)u(t)\right)dt, \\ & \qquad\qquad j = 1,2,\ldots,m_2, \\ & (y,v) \in \mathbf{G}(\omega). \end{array} \tag{43}$$

Assuming that the event space Ω is discrete and finite with K realizations, the discretization of $q_j(\cdot,\cdot,\omega)$ and $\mathbf{G}(\omega)$ are given by

$$q_j^{(k)}(y,v) \triangleq \frac{1}{2}\int_0^T \left(y(t)^{\mathsf{T}} \mathscr{Q}_j^{(k)}(t)y(t) + v(t)^{\mathsf{T}} \mathscr{S}_j^{(k)}(t)v(t)\right)dt, \text{ for } j = 0,1,\ldots,m_2,$$

and

$$\mathbf{G}^{(k)} \triangleq \left\{ (y,v) \in L_2^{m_2}[0,T] \times L_2^{p_2}[0,T] : \begin{array}{l} y \text{ is absolutely continuous on } [0,T], \\ \dot{y} \in L_2^{m_2}[0,T], y(0) = 0, \\ \dot{y}(t) = W^{(k)}(t)y(t) + D^{(k)}(t)v(t) \\ \text{for all } t \in [0,T] \end{array} \right\},$$

for $k = 1, 2, \ldots, K$, where $W^{(k)}(t)$ (respectively, $D^{(k)}(t), \mathscr{Q}_j^{(k)}(t), \mathscr{S}_j^{(k)}(t)$) is related to $W(t,\omega)$ (respectively, $D(t,\omega)$, $\mathscr{Q}_j(t,\omega), \mathscr{S}_j(t,\omega)$) in the same way as $w_j^{(k)}$ is related to $w_j(\omega)$ in Section 3. Given this, we have $\mathbb{E}[\rho(y,v,\omega)] = \sum_{k=1}^{K} q_0^{(k)}(y,v)$, and the discretization of (43) is:

$$\begin{aligned}
\min \quad & q_0^{(k)}(y,v) \\
\text{s.t.} \quad & q_j^{(k)}(y,v) + \int_0^T \left[w_j^{(k)^\top} y(t) + w_j^{(k)^\top} v(t) \right] dt \le h_j^{(k)} - \int_0^T \left[t_j^{(k)^\top} x(t) + t_j^{(k)^\top} u(t) \right] dt, \\
& \qquad\qquad\qquad\qquad\qquad\qquad j = 1, 2, \ldots, m_2, \\
& (y,v) \in \mathbf{G}^{(k)}.
\end{aligned} \tag{44}$$

Problems (42) and (44) are already on the form of Problems (6) and (8), respectively. In this context, the following logarithmic-barrier problem is the counterpart of Problem (14).

$$\begin{aligned}
\min \quad & f(x,u,\mu) = p_0 + \sum_{k=1}^{K} q_0^{(k)}(y,v) - \mu \left(\sum_{i=1}^{m_1} \ln(s_i(x,u)) + \sum_{j=1}^{m_2} \ln z_j^{(k)} \right) \\
\text{s.t.} \quad & \frac{1}{2} \int_0^T \left(x^\top(t) \mathscr{Q}_i(t) x(t) + u^\top(t) \mathscr{R}_i(t) u(t) \right) dt + s_i + b_i = 0, \ 1 \le i \le m_1, \\
& q_j^{(k)}(y,v) + \int_0^T \left(w_j^{(k)^\top} y(t) + w_j^{(k)^\top} v(t) \right) dt + z_j^{(k)} = h_j^{(k)} - \int_0^T \left(t_j^{(k)^\top} x(t) + t_j^{(k)^\top} u(t) \right) dt, \\
& \qquad\qquad\qquad\qquad\qquad\qquad 1 \le j \le m_2, 1 \le k \le K, \\
& (x,u) \in \mathbf{G}, (y,v) \in \mathbf{G}^{(k)}, \\
& s, z_j^{(k)} > 0, \ 1 \le i \le m_1, \ 1 \le j \le m_2, \ 1 \le k \le K,
\end{aligned}$$

where $\mu > 0$ is a barrier parameter. Here, $s_i \triangleq -p_i(x(t), u(t)), i = 1, 2, \ldots, m_1$, and

$$\begin{aligned}
z_j^{(k)} \ \triangleq \ & h_j^{(k)} - q_j^{(k)}(y(t), v(t)) - \int_0^T \left(w_j^{(k)^\top} y(t) + w_j^{(k)^\top} v(t) \right) dt \\
& - \int_0^T \left(t_j^{(k)^\top} x(t) + t_j^{(k)^\top} u(t) \right) dt, \ j = 1, \ldots, m_2, \ k = 1, \ldots, K.
\end{aligned}$$

Now, we compute the first and second derivatives of $p(x,u,\mu)$ based on the results in Section 4.

For the sake of simplicity, we drop the index time t and let $\alpha \triangleq (x,u) \in \mathcal{F}_1 \triangleq \{(x,u) \in \mathbf{G} : \ s = -p(x,u) > 0\}$. The first Fréchet derivative of p with respect to α is

$$\mathcal{D}_\alpha p(\mu, \alpha) = \begin{bmatrix} \mathscr{P}_0 x + \mu \sum_{i=1}^{m_1} \dfrac{\mathscr{P}_i x}{s_i(\alpha)} - \sum_{k=1}^{K} \sum_{j=1}^{m_2} \lambda_j(x) t_j^{(k)} \\ \mathscr{R}_0 u + \sum_{i=1}^{m_1} \dfrac{\mathscr{R}_i u}{s_i(\alpha)} - \sum_{k=1}^{K} \sum_{j=1}^{m_2} \lambda_j(u) t_j^{(k)} \end{bmatrix}, \tag{45}$$

or equivalently

$$\mathcal{D}_\alpha p(\mu,\alpha) = \begin{bmatrix} \mathscr{P}_0 x + \mu \mathscr{P}^\dagger[x] s^{-1}(\alpha) - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} \lambda(x) \\ \mathscr{R}_0 u + \mu \mathscr{R}^\dagger[u] s^{-1}(\alpha) - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} \lambda(u) \end{bmatrix}. \tag{46}$$

Next, for any $(\xi,\zeta) \in \mathbf{H}$, the second Fréchet derivative of p with respect to α is

$$\mathcal{D}^2_{\alpha\alpha} p(\mu,\alpha)(\xi,\zeta) = \begin{bmatrix} \mathscr{P}_0\xi + \mu \sum_{i=1}^{m_1} \left(\frac{\mathscr{P}_i \xi}{s_i(\alpha)} + \frac{l_i \mathscr{P}_i x}{s_i^2(\alpha)} \right) - \sum_{k=1}^{K} \sum_{j=1}^{m_2} \mathcal{D}_\alpha\left(\lambda_j^{(k)}(x)\right)(\xi) \left\langle t_j^{(k)}, \xi \right\rangle \\ \mathscr{R}_0\zeta + \mu \sum_{i=1}^{m_1} \left(\frac{\mathscr{R}_i \zeta}{s_i(\alpha)} + \frac{l_i \mathscr{R}_i u}{s_i^2(\alpha)} \right) - \sum_{k=1}^{K} \sum_{j=1}^{m_2} \mathcal{D}_\alpha\left(\lambda_j^{(k)}(u)\right)(\zeta) \left\langle t_j^{(k)}, \zeta \right\rangle \end{bmatrix},$$

or equivalently

$$\mathcal{D}^2_{\alpha\alpha} p(\mu,\alpha)(\xi,\zeta) = \begin{bmatrix} \mathscr{P}_0\xi + \mu \left(\mathcal{P}^\dagger[x]\right) S^{-2} \left(\mathcal{P}[x]\right) l - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} \mathcal{J}_\alpha\left[\lambda^{(k)}(x)\right]\xi \\ \mathscr{R}_0\zeta + \mu \left(\mathcal{R}^\dagger[x]\right) S^{-2} \left(\mathcal{R}[x]\right) l - \sum_{k=1}^{K} \mathcal{T}^{(k)\dagger} \mathcal{J}_\alpha\left[\lambda^{(k)}(u)\right]\zeta \end{bmatrix},$$

where

$$l_i \triangleq \int_0^T \left(x^\top(t) \mathscr{P}_i \xi + u^\top \mathscr{R}_i \zeta \right) dt, \; i = 1,2,\ldots,m_1$$

and more compactly

$$l \triangleq \int_0^T \left(\mathscr{Q}[\xi]x(t) + \mathscr{R}[\zeta]u(t) \right) dt.$$

Using (17), we get

$$\mathcal{D}^2_{\alpha\alpha} p(\mu,\alpha)(\xi,\zeta) = \begin{bmatrix} \mathscr{P}_0\xi + \mu \left(\mathcal{P}^\dagger[x]\right) S^{-2} \left(\mathcal{P}[x]\right) l - \sum_{k=1}^{K} \mathcal{T}^{(k\dagger} L^{(k)} N^{(k)} L^{(k)} \mathcal{T}^{(k)} \xi \\ \mathscr{R}_0\zeta + \mu \left(\mathcal{R}^\dagger[x]\right) S^{-2} \left(\mathcal{R}[x]\right) l - \sum_{k=1}^{K} \mathcal{T}^{(k\dagger} L^{(k)} N^{(k)} L^{(k)} \mathcal{T}^{(k)} \zeta \end{bmatrix}, \tag{47}$$

where the matrices $L^{(k)}$ and $N^{(k)}, k = 1,2,\ldots,K$, are defined in (18).
The orthogonal complements of subspaces $\mathbf{G}$ and $\mathbf{G}^{(k)}, k = 1,2,\ldots,K$, are given by [11]

$$\begin{aligned}\mathbf{G}^{\perp} &= \left\{\begin{bmatrix}\dot{\varpi}+A^{\mathsf{T}}\varpi\\ B^{\mathsf{T}}\varpi\end{bmatrix}: \varpi \text{ is absolutely continuous on } [0,T],\ \varpi(T)=0 \right.\\ &\qquad\qquad\qquad\qquad \left. \text{and } \dot{\varpi}\in L_2^n[0,T]\right\},\\ \mathbf{G}^{(k)\perp} &= \left\{\begin{bmatrix}\dot{\varrho}+W^{(k)^{\mathsf{T}}}\varrho\\ D^{(k)^{\mathsf{T}}}\varrho\end{bmatrix}: \varrho \text{ is absolutely continuous on } [0,T],\ \varrho(T)=0 \right.\\ &\qquad\qquad\qquad\qquad \left. \text{and } \dot{\varrho}\in L_2^n[0,T]\right\}.\end{aligned} \tag{48}$$

From (32), the Newton direction at the point $\alpha=(x,u)\in\mathbf{H}$ is of the form:

$$\begin{aligned}\triangle\alpha &= -\{\mathcal{D}^2_{\alpha\alpha}p(\mu,\alpha)\}^{-1}\left(\mathcal{D}_{\alpha}p(\mu,\alpha)-\beta^*(\alpha)\right),\\ \beta^*(\alpha) &\in \mathbf{G}^{\perp}.\end{aligned} \tag{49}$$

If $\triangle\alpha=\begin{bmatrix}\xi\\ \zeta\end{bmatrix}\in\mathbf{G}$, then by using (47) and (48) we have

$$\mathcal{D}_{\alpha}p(\mu,\alpha)-\beta^*(\alpha)=\mathcal{D}^2_{\alpha\alpha}p(\mu,\alpha)\triangle\alpha,\ \ \triangle\alpha\in\mathbf{G}.$$

Let $\mathcal{D}_{\alpha}p(\mu,\alpha)=\begin{bmatrix}\bar{x}\\ \bar{u}\end{bmatrix}$, then this is equivalent to

$$\begin{aligned}\dot{\varpi}+A^{\mathsf{T}}\varpi-\bar{x} &= \mathscr{K}_2(\mu,\alpha)\xi+\mathcal{P}^{\dagger}[x]S^{-2}(\alpha)l,\\ B^{\mathsf{T}}\varpi-\bar{u} &= \mathscr{K}_2(\mu,\alpha)\zeta+\mathcal{R}^{\dagger}[x]S^{-2}(\alpha)l,\end{aligned} \tag{50}$$

where

$$\mathscr{K}_1(\mu,\alpha)\triangleq\mathscr{P}_0+\mu\mathscr{P}^{\dagger}[\xi](s^{-1}(\alpha))-\sum_{k=1}^{K}\mathcal{T}^{(k\dagger}L^{(k)}N^{(k)}L^{(k)}\mathcal{T}^{(k)},$$

$$\mathscr{K}_2(\mu,\alpha)\triangleq\mathscr{R}_0+\mu\mathscr{R}^{\dagger}[\zeta](s^{-1}(\alpha))-\sum_{k=1}^{K}\mathcal{T}^{(k\dagger}L^{(k)}N^{(k)}L^{(k)}\mathcal{T}^{(k)}.$$

We also have

$$\dot{\xi}=A\xi+B\zeta,\qquad \xi(0)=0,\qquad \varpi(T)=0. \tag{51}$$

Now we need to solve (50)-(51) with respect to ξ and ζ. This problem is somehow similar to the one arising in connection with the standard linear-quadratic control problem (see[12]); the difference is that the constants l_i, $i=1,2,\ldots,m_1$, are unknown.

One can solve (50)-(51) by the numerical integration of Riccati equation that admits unique solution in $[0, T]$. To see this, we find ϖ in the form

$$\varpi(t) = K(t)\xi(t) + \gamma(t),$$

and substitute $\varpi(t)$ and (51) in both equations of (50) to obtain

$$\begin{aligned}&\dot{K} + KA + A^{\mathsf{T}}K + KB(t)\mathscr{K}_2^{-1}(\mu,\alpha)B^{\mathsf{T}}(t)K - \mathscr{K}_1(\mu,\alpha) = 0,\ K(T) = 0,\\ &\dot{\gamma}(t) = -\left(A^T + KB(t)\mathscr{K}_2(\mu,\alpha)^{-1}B^{\mathsf{T}}(t)K\right)\gamma + \bar{x} + \mathscr{P}^{\dagger}[x]S^{-2}(\alpha)l -\\ &B\left(\mathscr{K}_2^{-1}(\mu,\alpha)\bar{u} + \mathscr{K}_2^{-1}(\mu,\alpha)\mathscr{R}^{\dagger}[u]S^{-2}(\alpha)l\right), \gamma(T) = 0.\end{aligned} \tag{52}$$

The first equation in System (52) is a matrix Ricatti differential equation which admits unique solution in $[0, T]$ under natural constraints on A and B, and the second equation is a fundamental matrix for a linear time-dependent system. This brings us to solve an $m_1 \times m_1$ system of linear algebraic equation, and means that System (49) can be efficiently solved.

9. Conclusion

In this paper, we have studied the two-stage stochastic convex quadratic programming problem with quadratic constraints in Hilbert space and developed a logarithmic barrier decomposition interior point algorithm for solving this class of optimization problems. One of the chief attractions of this paper is that it explicitly computes the expressions for the derivatives of the recourse function and completely identifies the barrier parameters for the corresponding self-concordant family. Based on a self-concordance analysis, we have analyzed the proposed algorithm and have found that, given m_1 quadratic constraints in the first-stage problem, m_2 quadratic constraints in the second-stage problem, and K number of realizations, we need at most $\mathscr{O}((m_1 + Km_2)^{1/2}\ln(\mu^0/\epsilon))$ Newton iterations in short-step version of the algorithm to follow the first-stage central path from a starting value of the barrier μ^0 to a terminating value ϵ, and need at most $\mathscr{O}(m_1 + Km_2\ln(\mu^0/\epsilon))$ Newton iterations in the long-step version class of the algorithm to follow the first-stage central path. These complexity results coincide with the best known ones for the finite-dimensional case, confirming that the time execution of the proposed algorithm is independent of the underlying Hilbert space. As an example, we have considered an application of these results to an important stochastic control problem and have shown that the corresponding infinite-dimensional system can be obtained to find the Newton-type search direction. This study extends the work of Alzalg [4] for infinite-dimensional stochastic linear programming to infinite-dimensional stochastic quadratically constrained convex quadratic programming.

Conflict of interest. The author declares that they have no conflict of interest.

Data Availability. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] B. Alzalg, *Decomposition-based interior point methods for stochastic quadratic second-order cone programming*, Appl. Math. Comput. **249** (2014), 1–18.
https://doi.org/10.1016/j.amc.2014.10.015.

[2] ______, *Homogeneous self-dual algorithms for stochastic second-order cone programming*, J. Optim. Theory Appl. **163** (2014), no. 1, 148–164.
https://doi.org/10.1007/s10957-013-0428-z.

[3] ______, *Volumetric barrier decomposition algorithms for stochastic quadratic second-order cone programming*, Appl. Math. Comput. **265** (2015), 494–508.
https://doi.org/10.1016/j.amc.2015.05.014.

[4] ______, *Logarithmic-barrier decomposition interior-point methods for stochastic linear optimization in a Hilbert space*, Numer. Funct. Anal. Optim. **41** (2020), no. 8, 901–928.
https://doi.org/10.1080/01630563.2019.1709499.

[5] B. Alzalg and K.A. Ariyawansa, *Logarithmic barrier decomposition-based interior point methods for stochastic symmetric programming*, J. Math. Anal. Appl. **409** (2014), 973–995.
https://doi.org/10.1016/j.jmaa.2013.07.075.

[6] B. Alzalg, K. Badarneh, and A. Ababneh, *An infeasible interior-point algorithm for stochastic second-order cone optimization*, J. Optim. Theory Appl. **181** (2019), no. 1, 324–346.
https://doi.org/10.1007/s10957-018-1445-8.

[7] B. Alzalg, A. Gafour, and L. Alzaleq, *Volumetric barrier cutting plane algorithms for stochastic linear semi-infinite optimization*, IEEE Access **8** (2019), 4995–5008.
https://doi.org/10.1109/ACCESS.2019.2962840.

[8] B. Alzalg and M. Pirhaji, *Primal-dual path-following algorithms for circular programming*, Commun. Comb. Optim. **2** (2017), no. 2, 65–85.
https://doi.org/10.22049/cco.2017.25865.1051.

[9] K. Ariyawansa and Y. Zhu, *A class of polynomial volumetric barrier decomposition algorithms for stochastic semidefinite programming*, Math. Comput. **80** (2011), no. 275, 1639–1661.
https://doi.org/10.1090/s0025-5718-2010-02449-4.

[10] G.M. Cho, *Log-barrier method for two-stage quadratic stochastic programming*, Appl. Math. Comput. **164** (2005), no. 1, 45–69.
https://doi.org/10.1016/j.amc.2004.04.095.

[11] L. Faybusovich and J.B. Moore, *Infinite-dimensional quadratic optimization: interior-point methods and control applications*, Appl. Math. Optim. **36** (1997), no. 1, 43–66.

https://doi.org/10.1007/BF02683337.
[12] ———, *Long-step path-following algorithm for convex quadratic programming problems in a Hilbert space*, J. Optim. Theory Appl. **95** (1997), no. 3, 615–635.
https://doi.org/10.1023/A:1022626006554.
[13] S. Mehrotra and M. Gokhan Ozevin, *Decomposition based interior point methods for two-stage stochastic convex quadratic programs with recourse*, Oper. Res. **57** (2009), no. 4, 964–974.
https://doi.org/10.1287/opre.1080.0659.
[14] S. Mehrotra and M.G. Özevin, *Decomposition-based interior point methods for two-stage stochastic semidefinite programming*, SIAM J. Optim. **18** (2007), no. 1, 206–222.
https://doi.org/10.1137/050622067.
[15] S. Mehrotra and M.G. Ozevin, *Decomposition based interior point methods for two-stage stochastic convex quadratic programs with recourse*, Oper. Res. **57** (2009), no. 4, 964–974.
http://doi.org/10.1287/opre.1080.0659.
[16] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, 1994.
[17] J. Renegar, *Linear programming, complexity theory and elementary functional analysis*, Math. Program. **70** (1995), no. 1, 279–351.
https://doi.org/10.1007/BF01585941.
[18] ———, *A Mathematical View of Interior-Point Methods in Convex Optimization*, SIAM, 2001.
[19] C. Roos and J.P. Vial, *A polynomial method of approximate centers for linear programming*, Math. Program. **54** (1992), no. 1, 295–305.
https://doi.org/10.1007/BF01586056.
[20] G. Zhao, *A log-barrier method with Benders decomposition for solving two-stage stochastic linear programs*, Math. Program. **90** (2001), no. 3, 507–536.
https://doi.org/10.1007/PL00011433.

Appendices

A. Technical lemmas for the complexity proofs

In this appendix, we state without proofs some technical lemmas which were required to prove Theorems 3 and 4 in Section 7. The proofs of the following lemmas are based on Nestrov and Nemirovskii [16] for the finite-dimensional case and Faybusovich and Moore [12] for the infinite-dimensional case.
In the proof of Theorem 3, we made use of Theorem 3.1.1 of Nestrov and Nemirovskii [16] and Theorem 3.3 of Faybusovich and Moore [11], which are restated for our setting in the following lemmas.

Lemma 5. *Let* $\chi(\eta;\mu,\mu^+) \triangleq \left(\frac{1+\sqrt{m_2}}{2} + \frac{\sqrt{m_1+km_2}}{\kappa}\right)\ln\gamma^{-1}$. *Assume that* $\delta(x,\mu) < \kappa$ *and* $\mu^+ \triangleq \gamma\mu$ *satisfies* $\chi_k(\eta,\mu,\mu^+) \leq 1 - \delta(x,\mu)/\kappa$. *Then,* $\delta(\mu^+, x) < \kappa$.

Lemma 6. *Let* $\mu^+ = \gamma\mu$, *where* $\gamma = 1 - \delta(x,\mu)/\sqrt{m_1 + Km_2}$ *and* $\sigma \leq 0.1$. *Furthermore, let* $\beta = (2 - \sqrt{3})/2$. *If* $\delta(x,\mu) \leq \beta$, *then* $\delta(\mu^+, x) \leq 2\beta$.

In the proof of Theorem 4, we made use of the following lemma whose proof is similar to that of [20, Lemma 7].

Lemma 7. *For any* $\mu > 0$ *and* $x \in \mathcal{F}_1$, *we denote* $\tilde{\Delta}x \triangleq x - x(\mu)$ *and define*

$$\tilde{\delta} \triangleq \tilde{\delta}(x,\mu) \triangleq \sqrt{-\frac{1}{\mu}\mathcal{D}^2_{xx}\eta(x,\mu)\,(\tilde{\Delta}x^+, \tilde{\Delta}x^+)}. \tag{53}$$

For any $\mu > 0$ *and* $x \in \mathcal{F}_1$, *if* $\tilde{\delta} < 1$, *the following inequalities hold:*

$$\phi(x,\mu) \leq \mu\left(\frac{\tilde{\delta}}{1-\tilde{\delta}} + \ln(1-\tilde{\delta})\right), \quad \text{and} \quad \left|\phi'(x,\mu)\right| \leq -\sqrt{(m_1 + Km_2)}\,\ln((1-\tilde{\delta}).$$

Lemma 7 requires that $\tilde{\delta} < 1$. However, evaluation of $\tilde{\delta}$ explicitly may not be possible. Now we will see that $\tilde{\delta}$ is actually proportional to δ, which can be evaluated. The following lemma is due to [15] and its proof follows from Propositions 3.4 and 3.8 in [12] (see also [14, Lemma 5.5]).

Lemma 8. *For any* $\mu > 0$ *and* $x \in \mathcal{F}_1$, *we denote* $\tilde{\Delta}x \triangleq x = x(\mu)$ *and define* $\tilde{\delta} = \tilde{\delta}(x,\mu)$ *as in (53). If* $\delta < 1/6$, *then* $2\delta/3 \leq \tilde{\delta} \leq 2\delta$.