
TECHNICAL REPORT 2006-9

*A Class of Volumetric Barrier Decomposition
Algorithms for Stochastic Quadratic Programming*

K. A. Ariyawansa and Yuntao Zhu

August 2006

A Class of Volumetric Barrier Decomposition Algorithms for Stochastic Quadratic Programming

K. A. Ariyawansa* and Yuntao Zhu†

Abstract

Ariyawansa and Zhu have introduced a class of volumetric barrier decomposition algorithms [5] for solving two-stage stochastic semidefinite programs with recourse (SSDPs) [4]. In this paper we utilize their work for SSDPs to derive a class of volumetric barrier decomposition algorithms for solving two-stage stochastic quadratic programs with recourse and to establish polynomial complexity of certain members of the class of algorithms.

AMS Subject Classifications: 90C15, 90C22, 90C20, 90C51

Keywords: Quadratic Programming, Stochastic Programming, Volumetric Barrier, Self-Concordance, Decomposition

1 Introduction

Let $\mathbb{R}^{m \times n}$ and $\mathbb{R}^{n \vee n}$ denote the vector spaces of real $m \times n$ matrices and real symmetric $n \times n$ matrices respectively. For $U, V \in \mathbb{R}^{n \vee n}$ we write $U \succeq 0$ ($U \succ 0$) to mean that U is positive semidefinite (positive definite) and $U \succeq V$ or $V \preceq U$ to mean that $U - V \succeq 0$. For $U, V \in \mathbb{R}^{m \times n}$ we write $U \bullet V := \text{trace}(U^T V)$ to denote the Frobenius inner product between U and V . Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{k \times l}$ respectively. Then we define the Kronecker product $A \otimes B \in \mathbb{R}^{mk \times nl}$ of A and B as the matrix whose (i, j) block is $a_{ij}B$ for $i = 1, 2, \dots, m, j = 1, 2, \dots, n$. Given $U_i \in \mathbb{R}^{n_i \times n_i}$ for $i = 1, 2, \dots, n$ we use $\text{diag}(U_1, U_2, \dots, U_n)$ to denote the matrix in $\mathbb{R}^{(\sum_{i=1}^n n_i) \times (\sum_{i=1}^n n_i)}$ with U_1, U_2, \dots, U_n on the diagonal and zeros elsewhere. For any matrix $A \in \mathbb{R}^{m \times n}$, we use $[A]_i \in \mathbb{R}^m$ to denote the i -th column of A for $i = 1, 2, \dots, n$, and we use $\text{vec}(A) \in \mathbb{R}^{mn}$ to denote the vector formed by “stacking” the columns of A one atop another in the natural order.

A (two-stage) stochastic quadratic program (with recourse) (SQP) is defined based on deterministic data $C \in \mathbb{R}^{n_1 \vee n_1}$, $C \succeq 0$, $c \in \mathbb{R}^{n_1}$, $A \in \mathbb{R}^{m_1 \times n_1}$, and $b \in \mathbb{R}^{m_1}$; and random data $H \in \mathbb{R}^{n_2 \vee n_2}$, $H \succeq 0$, $d \in \mathbb{R}^{n_2}$, $T \in \mathbb{R}^{m_2 \times n_1}$, $W \in \mathbb{R}^{m_2 \times n_2}$, and $h \in \mathbb{R}^{m_2}$ whose realizations depend on an underlying outcome ω in an event space Ω with a known probability function P . Given this data, an SQP with recourse is

$$\begin{aligned} & \text{minimize} && \frac{1}{2}x^T Cx + c^T x + \mathbb{E}[Q(x, \omega)] \\ & \text{subject to} && Ax \leq b \end{aligned} \tag{1}$$

*Department of Mathematics, Washington State University, Pullman, WA 99164-3113, USA. (ari@wsu.edu). The work of this author was supported in part by the U.S. Army Research Office under Grant DAAD 19-00-1-0465.

†Department of Mathematics, Washington State University, Pullman, WA 99164-3113, USA. (zhuyt@wsu.edu). The material in this paper is part of the doctoral dissertation [15] of this author completed at Washington State University.

where $x \in \mathbb{R}^{n_1}$ is the first-stage decision variable, $Q(x, \omega)$ is the minimum of the problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2}y^\top H(\omega)y + d(\omega)^\top y \\ & \text{subject to} && T(\omega)x + W(\omega)y \leq h(\omega) \end{aligned} \quad (2)$$

where $y \in \mathbb{R}^{n_2}$ is the second-stage variable, and

$$\mathbb{E}[Q(x, \omega)] := \int_{\Omega} Q(x, \omega) P(d\omega). \quad (3)$$

We note that solving the SQP (1, 2, 3) is equivalent to solving [12, Problem P] due to the duality theory for quadratic programming. (See also [10, Chapter 8, Section 2].)

We now examine the SQP (1, 2, 3) when the event space Ω is finite. Let $\{(H^{(k)}, d^{(k)}, T^{(k)}, W^{(k)}, h^{(k)}) : k = 1, 2, \dots, K\}$ be the possible values of the random variables $(H(\omega), d(\omega), T(\omega), W(\omega), D(\omega))$ and let $p_k := P((H(\omega), d(\omega), T(\omega), W(\omega), D(\omega)) = (H^{(k)}, d^{(k)}, T^{(k)}, W^{(k)}, h^{(k)}))$ be the associated probability for $k = 1, 2, \dots, K$. Then Problem (1, 2, 3) becomes

$$\begin{aligned} & \text{minimize} && \frac{1}{2}x^\top Cx + c^\top x + \sum_{k=1}^k p_k Q^{(k)}(x) \\ & \text{subject to} && Ax \leq b \end{aligned} \quad (4)$$

where $x \in \mathbb{R}^{n_1}$ is the first-stage variable, and $Q^{(k)}(x)$ is the minimum of the problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2}(y^{(k)})^\top H^{(k)}y^{(k)} + (d^{(k)})^\top y^{(k)} \\ & \text{subject to} && T^{(k)}x + W^{(k)}y^{(k)} \leq h^{(k)} \end{aligned} \quad (5)$$

where $y^{(k)} \in \mathbb{R}^{n_2}$ is the second-stage variable, for $k = 1, 2, \dots, K$.

In order to conveniently make use of the results in [5], we rewrite Problem (4, 5) in the following equivalent form by multiplying each inequality in the constraints by (-1) and then replacing each vector and each matrix in the constraints by its negative respectively:

$$\begin{aligned} & \text{minimize} && \frac{1}{2}x^\top Cx + c^\top x + \sum_{k=1}^K p_k Q^{(k)}(x) \\ & \text{subject to} && Ax \geq b \end{aligned} \quad (6)$$

where $x \in \mathbb{R}^{n_1}$ is the first-stage variable, and $Q^{(k)}(x)$ is the minimum of the problem

$$\begin{aligned} & \text{minimize} && \frac{1}{2}(y^{(k)})^\top H^{(k)}y^{(k)} + (d^{(k)})^\top y^{(k)} \\ & \text{subject to} && T^{(k)}x + W^{(k)}y^{(k)} \geq h^{(k)} \end{aligned} \quad (7)$$

where $y^{(k)} \in \mathbb{R}^{n_2}$ is the second-stage variable, for $k = 1, 2, \dots, K$.

In the rest of this paper our attention will be only on Problem (6, 7), and from now on when we use the term stochastic quadratic program (SQP) in this paper we mean Problem (6, 7).

The case of Problem (6, 7) with $C := 0$ and $H^{(k)} := 0$ for $k = 1, 2, \dots, K$ is a stochastic linear program (SLP). In [14], Zhao presented a class of decomposition algorithms based on a logarithmic barrier for SLPs and proved polynomial complexity results for certain member of the class. Cho [6] extended the work of Zhao [14] to the case of SQPs. The algorithms of Cho [6] are also based on logarithmic barriers.

In [4], Ariyawansa and Zhu defined a new class of optimization problems termed (two-stage) stochastic semidefinite programs (with recourse) (SSDPs) extending SLPs, and in [5] they extended

the class of algorithms of Zhao [14] to SSDPs but used volumetric barriers [13] in place of logarithmic barriers. The use of the technically more complicated volumetric barrier instead of the logarithmic barrier was motivated by computational results in [7] which showed that in the case of certain cutting plane algorithms [9] for SLPs, algorithms based on volumetric barrier perform better than corresponding algorithms based on the logarithmic barrier.

The authors know of no work based on volumetric barrier analogous to the work of Zhao [14] for SLPs and SQPs. The purpose of this paper is to present a class of such algorithms for SQPs (and hence for SLPs as well). The present paper makes crucial use of results obtained in [5].

In [5] we considered the following problem:

$$\begin{aligned} & \text{minimize} && b^\top y + \sum_{k=1}^K p_k Q^{(k)}(y) \\ & \text{subject to} && \sum_{i=1}^{m_1} y_i A_i - C \succeq 0, \end{aligned} \tag{8}$$

where for $k = 1, 2, \dots, K$, $Q^{(k)}(y)$ is the minimum of

$$\begin{aligned} & \text{minimize} && (d^{(k)})^\top x^{(k)} \\ & \text{subject to} && \sum_{i=1}^{m_1} y_i W_i^{(k)} + \sum_{i=1}^{m_2} x_i^{(k)} T_i^{(k)} - D^{(k)} \succeq 0. \end{aligned} \tag{9}$$

Problem (8, 9) is an SSDP (see [4, 5]). In [5] we derived a class of volumetric barrier decomposition algorithms for solving (8, 9).

Now we demonstrate how the constraints in (6, 7) can be converted in the form of constraints in (8, 9). We make the following assignments:

$$\begin{aligned} A_i & := \text{diag}([A]_i), & i = 1, 2, \dots, n_1; \\ C & := \text{diag}(b); \\ T_i^{(k)} & := \text{diag}([T^{(k)}]_i), & i = 1, 2, \dots, n_1, k = 1, 2, \dots, K; \\ W_i^{(k)} & := \text{diag}([W^{(k)}]_i), & i = 1, 2, \dots, n_2, k = 1, 2, \dots, K; \\ D^{(k)} & := \text{diag}(h^{(k)}), & k = 1, 2, \dots, K. \end{aligned} \tag{10}$$

With these assignments the linear constraints in Problem (6, 7) are converted into equivalent semidefinite constraints. Hence many results we obtained in [5] can be utilized in our derivation of a class of volumetric barrier decomposition algorithms for the SQP (6, 7).

The paper is organized as follows. In the next section we introduce a volumetric barrier for the SQP (6, 7). In §3 we show that the set of barrier functions for positive values of the barrier parameter comprises a self-concordant family [11]. Based on this property a class of volumetric barrier decomposition algorithms is presented in §4. A convergence and complexity analysis of this class of certain members of this class of algorithms is presented in §5.

2 A Volumetric Barrier for SQPs

In this section we formulate a volumetric barrier for SQPs and obtain expressions for the derivatives required in the rest of the paper.

2.1 Formulation

In order to define the volumetric barrier problem for the SQP (6, 7), we need to make some assumptions. First we define

$$\begin{aligned}\mathcal{F}_1 &:= \{x \in \mathbb{R}^{n_1} : S_1(x) := Ax - b > 0\}; \\ \mathcal{F}_p^{(k)}(x) &:= \{y^{(k)} \in \mathbb{R}^{n_2} : S_2^{(k)}(x, y^{(k)}) := ((T^{(k)})x + W^{(k)})y - h^{(k)} > 0\}; \\ \mathcal{F}_2 &:= \{x \in \mathbb{R}^{n_1} : \mathcal{F}_p^{(k)}(x) \neq \emptyset, k = 1, 2, \dots, K\}; \\ \mathcal{F}_0 &:= \mathcal{F}_1 \cap \mathcal{F}_2.\end{aligned}$$

Then we make

Assumption 1. The set \mathcal{F}_0 is nonempty.

The set \mathcal{F}_1 is nonempty under Assumption 1. The logarithmic barrier [11] for \mathcal{F}_1 is the function $f_1 : \mathcal{F}_1 \rightarrow \mathbb{R}$ defined by

$$f_1(x) := - \sum_{i=1}^{m_1} \ln [S_1(x)]_i, \quad \forall x \in \mathcal{F}_1,$$

and the volumetric barrier [13, 11] for \mathcal{F}_1 is the function $V_1 : \mathcal{F}_1 \rightarrow \mathbb{R}$ defined by

$$V_1(x) := \frac{1}{2} \ln \det(\nabla_{xx}^2 f_1(x)), \quad \forall x \in \mathcal{F}_1.$$

Also under Assumption 1, \mathcal{F}_2 is nonempty and for $x \in \mathcal{F}_2$, $\mathcal{F}_p^{(k)}(x)$ is nonempty for $k = 1, 2, \dots, K$. The logarithmic barrier [11] for $\mathcal{F}_p^{(k)}(x)$ is the function $f_2^{(k)} : \mathcal{F}_p^{(k)}(x) \rightarrow \mathbb{R}$ defined by

$$f_2^{(k)}(x, y^{(k)}) := - \sum_{i=1}^{m_2} \ln [S_2^{(k)}(x, y^{(k)})]_i, \quad \forall y^{(k)} \in \mathcal{F}_p^{(k)}(x), x \in \mathcal{F}_2,$$

and the volumetric barrier [13, 11] for $\mathcal{F}_p^{(k)}(x)$ is the function $V_2^{(k)} : \mathcal{F}_p^{(k)}(x) \rightarrow \mathbb{R}$ defined by

$$V_2^{(k)}(x, y^{(k)}) := \frac{1}{2} \ln \det(\nabla_{y^{(k)}y^{(k)}}^2 f_2^{(k)}(x, y^{(k)})), \quad \forall y^{(k)} \in \mathcal{F}_p^{(k)}(x), x \in \mathcal{F}_2.$$

We note that for any $v \in \mathbb{R}^n$ such that $v > 0$, we have $\sum_{i=1}^n \ln v_i = \ln \det(\text{diag}(v))$. So under the assignment (10), the logarithmic barriers f_1 and $f_2^{(k)}$ for $k = 1, 2, \dots, K$, and the volumetric barriers V_1 and $V_2^{(k)}$ for $k = 1, 2, \dots, K$ we just defined above are equivalent to the corresponding logarithmic and volumetric barriers we defined in [5].

Next we make

Assumption 2. For each $x \in \mathcal{F}_0$ and for $k = 1, 2, \dots, K$, Problem (7) has a nonempty isolated compact set of minimizers.

We now define the volumetric barrier problem for the SQP (6, 7) as

$$\text{minimize } \eta(\mu, x) := \frac{1}{2}x^\top Cx + c^\top x + \sum_{k=1}^K p_k \rho_k(\mu, x) + \mu c_1 V_1(x) \quad (11)$$

where for $k = 1, 2, \dots, K$ and $x \in \mathcal{F}_0$, $\rho_k(\mu, x)$ is the minimum of

$$\text{minimize } \frac{1}{2}(y^{(k)})^\top H^{(k)}y^{(k)} + (d^{(k)})^\top y^{(k)} + \mu c_2 V_2^{(k)}(x, y^{(k)}) \quad (12)$$

Here $c_1 > 0$ and $c_2 > 0$ are constants whose values will be defined more precisely in the sequel and $\mu > 0$ is the barrier parameter.

We will now show that (12) has a unique minimizer for each $x \in \mathcal{F}_0$ and for $k = 1, 2, \dots, K$ by utilizing:

Theorem 1 (Fiacco and McCormick [8, Theorem 8]). *Consider the inequality constrained problem*

$$\begin{aligned} & \text{minimize } f(x) \\ & \text{subject to } g_i(x) \geq 0, \quad i = 1, 2, \dots, m, \end{aligned} \quad (13)$$

where the functions $f, g_1, \dots, g_m : \mathbb{R}^n \rightarrow \mathbb{R}$ are continuous. Let I be a scalar-valued function of x with the following two properties: $I(x)$ is continuous in the region $R^0 := \{x : g_i(x) > 0, i = 1, 2, \dots, m\}$, which is assumed to be nonempty; if $\{x_k\}$ is any infinite sequence of points in R^0 converging to x_B such that $g_i(x_B) = 0$ for at least one i , then $\lim_{k \rightarrow \infty} I(x_k) = +\infty$. Let s be a scalar-valued function of the single variable r with the following two properties: if $r_1 > r_2 > 0$, then $s(r_1) > s(r_2) > 0$; if $\{r_k\}$ is an infinite sequence of points such that $\lim_{k \rightarrow \infty} r_k = 0$, then $\lim_{k \rightarrow \infty} s(r_k) = 0$. Let $U : R^0 \times \mathbb{R}^+ \rightarrow \mathbb{R}$ be defined by $U(x, r) := f(x) + s(r)I(x)$. If (13) has a nonempty, isolated compact set of local minimizers and $\{r_k\}$ is a strictly decreasing infinite sequence, then the unconstrained local minimizers of $U(\cdot, r_k)$ exist for r_k small.

Lemma 1. *If Assumptions 1 and 2 hold, then for each $x \in \mathcal{F}_0$ and $k = 1, 2, \dots, K$, the Problem (12) has a unique minimizer for μ small.*

Proof. By Theorem 1, local minimizers of (12) exist for each $x \in \mathcal{F}_0$ and $k = 1, 2, \dots, K$ for μ small. The uniqueness of the minimizer follows from the fact that $V_2^{(k)}$ is strictly convex. \square

By Lemma 1, the Problem (11, 12) is well-defined, and its feasible set is \mathcal{F}_0 .

2.2 Expressions for partial derivatives of η with respect to y

In order to compute the derivatives of η we need the derivatives of ρ_k , $k = 1, 2, \dots, K$, which in turn require the derivatives of $V_2^{(k)}$ and $f_2^{(k)}$ for each $k = 1, 2, \dots, K$. We need the following two theorems in our calculation of derivatives of $\rho_k(\mu, x)$ and we will show that the results we obtained in [5] can be utilized here to calculate the derivatives of $V_2^{(k)}$ and $f_2^{(k)}$ for $k = 1, 2, \dots, K$.

Theorem 2 (Clairaut's Theorem). *Suppose that f is a function of two variables and it's defined on a disk D that contains the point (a, b) . Let*

$$f_{xy} = \frac{\partial^2 f}{\partial y \partial x},$$

and

$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y}.$$

If the functions f_{xy} and f_{yx} are both continuous on this disk, then

$$f_{xy}(a, b) = f_{yx}(a, b).$$

Theorem 3 (An extension of Clairaut's Theorem). *Let*

$$f_{xxy} = \frac{\partial^3 f}{\partial y \partial x \partial x},$$

$$f_{xyx} = \frac{\partial^3 f}{\partial x \partial y \partial x},$$

and

$$f_{yxx} = \frac{\partial^3 f}{\partial x \partial x \partial y}.$$

If the functions f_{xyx} , f_{yxx} and f_{xxy} are all continuous, then

$$f_{xxy} = f_{xyx} = f_{yxx}.$$

Proof. If the third order derivatives are continuous then the second order ones are also continuous. So by Clairaut's Theorem we have

$$f_{xxy} = (f_x)_{xy} = (f_x)_{yx} = f_{xyx} = (f_{xy})_x = (f_{yx})_x = f_{yxx}. \quad \square$$

In general, we can extend Clairaut's Theorem to any order of mixed partial derivatives. The only requirement is that in each derivative we differentiate with respect to each variable the same number of times.

Now for the Problem (12) we have

$$\rho_k(\mu, x) = \frac{1}{2}(\bar{y}^{(k)})^\top H^{(k)} \bar{y}^{(k)} + (d^{(k)})^\top \bar{y}^{(k)} + \mu c_2 V_2^{(k)}(x, \bar{y}^{(k)}), \quad (14)$$

where $\bar{y}^{(k)}$ is the optimal solution for (12), for $k = 1, 2, \dots, K$. We notice that $\bar{y}^{(k)}$ is a function of x and the relationship between $\bar{y}^{(k)}$ and x is defined by

$$\frac{\partial}{\partial y^{(k)}} \left(\frac{1}{2}(y^{(k)})^\top H^{(k)} y^{(k)} + (d^{(k)})^\top y^{(k)} + \mu c_2 V_2^{(k)}(x, y^{(k)}) \right) \Big|_{y^{(k)} = \bar{y}^{(k)}} = 0. \quad (15)$$

Hence we have

$$\frac{\partial}{\partial \bar{y}^{(k)}} \left(\frac{1}{2}(\bar{y}^{(k)})^\top H^{(k)} \bar{y}^{(k)} + (d^{(k)})^\top \bar{y}^{(k)} + \mu c_2 V_2^{(k)}(x, \bar{y}^{(k)}) \right) = 0. \quad (16)$$

Proposition 1.

$$\frac{\partial^2}{\partial \bar{y}^{(k)} \partial x} V_2^{(k)}(x, \bar{y}^{(k)}) = 0.$$

Proof. By (16), we have

$$\frac{\partial}{\partial \bar{y}^{(k)}} V_2^{(k)}(x, \bar{y}^{(k)}) = (-d^{(k)} - H^{(k)} \bar{y}^{(k)}) / (\mu c_2).$$

Thus by Theorem 2, we have

$$\frac{\partial^2}{\partial \bar{y}^{(k)} \partial x} V_2(x, \bar{y}^{(k)}) = \frac{\partial^2}{\partial x \partial \bar{y}^{(k)}} V_2(x, \bar{y}^{(k)}) = 0. \quad \square$$

Proposition 2.

$$\frac{\partial^3}{\partial \bar{y}^{(k)} \partial x \partial x} V_2^{(k)}(x, \bar{y}^{(k)}) = 0.$$

Proof. By Theorem 3 and Proposition 1, the result follows. \square

Now we are ready to calculate the derivatives of ρ_k with respect to x . We have

$$\begin{aligned} \nabla_x \rho_k(\mu, x) &= \nabla_x \left(\frac{1}{2} (\bar{y}^{(k)})^\top H^{(k)} \bar{y}^{(k)} + (d^{(k)})^\top \bar{y}^{(k)} + \mu c_2 V_2^{(k)}(x, \bar{y}^{(k)}) \right) \\ &= \frac{\partial \rho_k(\mu, x)}{\partial \bar{y}^{(k)}} \cdot \frac{\partial \bar{y}^{(k)}}{\partial x} + \frac{\partial \rho_k(\mu, y)}{\partial x} \\ &= \frac{\partial \rho_k(\mu, x)}{\partial x} \quad (\text{by (16)}) \\ &= \mu c_2 \nabla_x V_2(x, \bar{y}^{(k)}), \end{aligned}$$

and

$$\begin{aligned} \nabla_{xx}^2 \rho_k(\mu, x) &= \nabla_x (\nabla_x \rho_k(\mu, x)) \\ &= \nabla_x (\mu c_2 \nabla_x V_2(x, \bar{y}^{(k)})) \\ &= \mu c_2 \frac{\partial}{\partial \bar{y}^{(k)}} (\nabla_x V_2(x, \bar{y}^{(k)})) \cdot \frac{\partial \bar{y}^{(k)}}{\partial x} + \mu c_2 \frac{\partial}{\partial x} (\nabla_x V_2(x, \bar{y}^{(k)})) \\ &= \mu c_2 \frac{\partial^2}{\partial \bar{y}^{(k)} \partial x} V_2(x, \bar{y}^{(k)}) \cdot \frac{\partial \bar{y}^{(k)}}{\partial x} + \mu c_2 \frac{\partial^2}{\partial x \partial x} V_2(x, \bar{y}^{(k)}) \\ &= \mu c_2 \nabla_{xx}^2 V_2(x, \bar{y}^{(k)}) \quad (\text{by Proposition (1)}). \end{aligned}$$

We also have

$$\begin{aligned} \nabla_{xxx}^3 \rho_k(\mu, x) &= \nabla_x (\nabla_{xx}^2 \rho_k(\mu, x)) \\ &= \nabla_x (\mu c_2 \nabla_{xx}^2 V_2(x, \bar{y}^{(k)})) \\ &= \mu c_2 \frac{\partial}{\partial \bar{y}^{(k)}} (\nabla_{xx}^2 V_2(x, \bar{y}^{(k)})) \cdot \frac{\partial \bar{y}^{(k)}}{\partial x} + \mu c_2 \frac{\partial}{\partial x} (\nabla_{xx}^2 V_2(x, \bar{y}^{(k)})) \\ &= \mu c_2 \frac{\partial^3}{\partial \bar{y}^{(k)} \partial x \partial x} V_2(x, \bar{y}^{(k)}) \cdot \frac{\partial \bar{y}^{(k)}}{\partial x} + \mu c_2 \frac{\partial^3}{\partial x \partial x \partial x} V_2(x, \bar{y}^{(k)}) \\ &= \mu c_2 \frac{\partial^3}{\partial x \partial x \partial \bar{y}^{(k)}} V_2(x, \bar{y}^{(k)}) \cdot \frac{\partial \bar{y}^{(k)}}{\partial x} + \mu c_2 \frac{\partial^3}{\partial x \partial x \partial x} V_2(x, \bar{y}^{(k)}) \\ &= \mu c_2 \nabla_{xxx}^3 V_2(x, \bar{y}^{(k)}) \quad (\text{by Proposition 2}). \end{aligned}$$

In summary we have

$$\begin{aligned} \nabla_x \rho_k(\mu, x) &= \mu c_2 \nabla_x V_2^{(k)}(x, \bar{y}^{(k)}), \\ \nabla_{xx}^2 \rho_k(\mu, x) &= \mu c_2 \nabla_{xx}^2 V_2^{(k)}(x, \bar{y}^{(k)}), \\ \nabla_{xxx}^3 \rho_k(\mu, x) &= \mu c_2 \nabla_{xxx}^3 V_2^{(k)}(x, \bar{y}^{(k)}), \end{aligned} \tag{17}$$

and

$$\begin{aligned} \nabla_x \eta(\mu, x) &= Cx + c + \mu c_1 \nabla_x V_1(x) + \sum_{k=1}^K \mu c_2 \nabla_x V_2^{(k)}(x, \bar{y}^{(k)}), \\ \nabla_{xx}^2 \eta(\mu, x) &= C + \mu c_1 \nabla_{xx}^2 V_1(x) + \sum_{k=1}^K \mu c_2 \nabla_{xx}^2 V_2^{(k)}(x, \bar{y}^{(k)}), \end{aligned} \tag{18}$$

Now we show how we can calculate $\nabla_x V_2^{(k)}(x, \bar{y}^{(k)})$ and $\nabla_{xx}^2 V_2^{(k)}(x, \bar{y}^{(k)})$ respectively. Some of these computations are lengthy and it is convenient to drop the superscript (k) . We do so when it does not lead to confusion.

First we make

Assumption 3. The matrix $W^{(k)}$ has full rank for $k = 1, 2, \dots, K$.

Let $\mathcal{W} \in \mathbb{R}^{m_2^2 \times n_2}$ be the matrix whose i -th column is $\text{vec}(W_i) \in \mathbb{R}^{m_2^2}$ for $i = 1, 2, \dots, n_2$. Then under the assignment (10) and results in [5, §3.2] the Hessian matrix $H := \nabla_{yy}^2 f_2(x, y)$ can be represented in the form:

$$H := \nabla_{yy}^2 f_2(x, y) = \mathcal{W}^\top [(S_2^{-1} \otimes S_2^{-1}) \mathcal{W}].$$

Note that by Assumption 3, H is positive definite. We have

$$\begin{aligned} \frac{\partial V_2(x, y)}{\partial x_i} &= -(\mathcal{W}H^{-1}\mathcal{W}^\top) \bullet (S_2^{-1}T_i S_2^{-1} \otimes_s S_2^{-1}) \\ &= -P \bullet (S_2^{-1/2}T_i S_2^{-1/2} \otimes_s I), \end{aligned}$$

for $i = 1, 2, \dots, n_1$, where

$$P = P(S_2) = (S_2^{-1/2} \otimes S_2^{-1/2}) \mathcal{W} (\mathcal{W}^\top (S_2^{-1} \otimes S_2^{-1}) \mathcal{W})^{-1} \mathcal{W}^\top (S_2^{-1/2} \otimes S_2^{-1/2})$$

is the orthogonal projection onto the range of $(S_2^{-1/2} \otimes S_2^{-1/2}) \mathcal{W}$.

We also have

$$\nabla_{xx}^2 V_2(x, y) = \frac{\partial^2}{\partial x \partial x} V_2(x, y) = 2Q^{xx} + R^{xx} - 2T^{xx},$$

where

$$\begin{aligned} Q_{i,j}^{xx} &= (\mathcal{W}H^{-1}\mathcal{W}^\top) \bullet (S_2^{-1}T_i S_2^{-1}T_j S_2^{-1} \otimes_s S_2^{-1}), \\ R_{i,j}^{xx} &= (\mathcal{W}H^{-1}\mathcal{W}^\top) \bullet (S_2^{-1}T_i S_2^{-1} \otimes_s S_2^{-1}T_j S_2^{-1}), \\ T_{i,j}^{xx} &= (\mathcal{W}H^{-1}\mathcal{W}^\top) \bullet (S_2^{-1}T_i S_2^{-1} \otimes_s S_2^{-1}) \mathcal{W}H^{-1}\mathcal{W}^\top (S_2^{-1}T_j S_2^{-1} \otimes_s S_2^{-1}). \end{aligned}$$

3 Characteristics of η : a self-concordant family

3.1 Self-Concordance of $\eta(\mu, \cdot)$

Definition 1 (Nesterov and Nemirovskii [11, Definition 2.1.1]). Let G be an open nonempty convex subset of \mathbb{R}^n , and let F be a \mathcal{C}^3 , convex mapping from G to \mathbb{R} . Then F is called α -self-concordant on G with the parameter $\alpha > 0$ if for every $x \in G$ and $\xi \in \mathbb{R}^n$, the following inequality holds:

$$|D^3 F(x)[\xi, \xi, \xi]| \leq 2\alpha^{-1/2} (D^2 F(x)[\xi, \xi])^{3/2}.$$

An α -self-concordant function F on G is called strongly α -self-concordant if F tends to infinity for any sequence approaching a boundary point of G .

We note that in the definition above the set G is assumed to be open. However, relative openness would be sufficient to apply the definition. See also [11, Item A, Page 57]. We now show that $\rho_k(\mu, \cdot)$ is μ -self-concordant on \mathcal{F}_0 , for $k = 1, 2, \dots, K$. It is clear that \mathcal{F}_0 is open.

Theorem 4. For any fixed $\mu > 0$ and $c_2 := 225 m_2^{3/2}$, $\rho_k(\mu, \cdot)$ is μ -self-concordant on \mathcal{F}_0 , for $k = 1, 2, \dots, K$.

Proof. This follows the proof in [5, Proof of Thoerem 4]. \square

Corollary 1. *For any fixed $\mu > 0$, $c_1 := 225\sqrt{m_1}$ and $c_2 := 225m_2^{3/2}$, $\eta(\mu, \cdot)$ is a μ -self-concordant function on \mathcal{F}_0 .*

Proof. Anstreicher [2, Theorem 5.1] has shown that $\mu c_1 V_1$ is μ -self-concordant on \mathcal{F}_1 . It is easy to verify that $(x^\top Cx + c^\top x + \mu c_1 V_1)$ is also μ -self-concordant on \mathcal{F}_1 . Indeed, for any $\xi \in \mathbb{R}^{n_1}$ we have

$$\begin{aligned} \nabla_{xxx}^3(x^\top Cx + c^\top x + \mu c_1 V_1)[\xi, \xi, \xi] &= \mu c_1 \nabla_{xxx}^3 V_1[\xi, \xi, \xi] \\ &\leq 2\mu^{-1/2}(\mu c_1 \nabla_{xx}^2 V_1[\xi, \xi])^{3/2} \\ &\leq 2\mu^{-1/2}((C + \mu c_1 \nabla_{xx}^2 V_1)[\xi, \xi])^{3/2} \\ &= 2\mu^{-1/2}(\nabla_{xx}^2(x^\top Cx + c^\top x + \mu c_1 V_1)[\xi, \xi])^{3/2} \end{aligned}$$

The corollary follows from [11, Proposition 2.1.1]. \square

From here onwards we fix the values of c_1 and c_2 as in Corollary 1.

3.2 Parameters of the Self-Concordant Family $\eta(\mu, \cdot)$

Definition 2 (Nesterov and Nemirovskii [11, Definition 3.1.1]). Let \mathbb{R}^{++} be the set of all positive real numbers. Let G be an open nonempty convex subset of \mathbb{R}^n . Let $\mu \in \mathbb{R}^{++}$ and let $F_\mu : \mathbb{R}^{++} \times G \rightarrow \mathbb{R}$ be a family of functions indexed by μ . Let $\alpha_1(\mu), \alpha_2(\mu), \alpha_3(\mu), \alpha_4(\mu)$ and $\alpha_5(\mu) : \mathbb{R}^{++} \rightarrow \mathbb{R}^{++}$ be continuously differentiable functions on μ . Then the family of functions $\{F_\mu\}_{\mu \in \mathbb{R}^{++}}$ is called strongly self-concordant with the parameters $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5$, if the following conditions hold:

- (i) F_μ is continuous on $\mathbb{R}^{++} \times G$, and for fixed $\mu \in \mathbb{R}^{++}$, F_μ is convex on G . F_μ has three partial derivatives on 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 $(\mu, x) \in \mathbb{R}^{++} \times G$ and any $\xi \in \mathbb{R}^n$,

$$\begin{aligned} |\{\nabla_x F_\mu(\mu, x)[\xi]\}' - \{\ln \alpha_3(\mu)\}' \nabla_x F_\mu(\mu, x)[\xi]| &\leq \alpha_4(\mu) \alpha_1(\mu)^{\frac{1}{2}} (\nabla_{xx}^2 F_\mu(\mu, x)[\xi, \xi])^{\frac{1}{2}}, \\ |\{\nabla_{xx}^2 F_\mu(\mu, x)[\xi, \xi]\}' - \{\ln \alpha_2(\mu)\}' \nabla_{xx}^2 F_\mu(\mu, x)[\xi, \xi]| &\leq 2\alpha_5(\mu) \nabla_{xx}^2 F_\mu(\mu, x)[\xi, \xi]. \end{aligned}$$

Theorem 5. *The parametric function $\eta(\mu, \cdot)$ is a strongly self-concordant family with the following parameters*

$$\alpha_1(\mu) = \mu, \alpha_2(\mu) = \alpha_3(\mu) = 1, \alpha_4(\mu) = \mu^{-1}((n_1 c_1 + n_2 c_2)(1 + K))^{1/2}, \alpha_5(\mu) = \frac{1}{\mu}.$$

By Corollary 1, in order to prove Theorem 5, we only need to show that the two inequalities in Definition 2 (iii) are satisfied by $\eta(\mu, \cdot)$. We first show the validity of the second inequality in Definition 2 (iii).

Lemma 2. *For any $\mu > 0$ and $x \in \mathcal{F}_0$, the following inequality holds:*

$$|\nabla_{xx}^2 \eta'(\mu, x)[\xi, \xi]| \leq \frac{1}{\mu} \nabla_{xx}^2 \eta(\mu, x)[\xi, \xi], \quad \forall \xi \in \mathbb{R}^{n_1}.$$

Proof. Differentiating (18) with respect to μ and using Proposition 1 and Proposition 2 we obtain

$$\begin{aligned}
\nabla_{xx}^2 \eta'(\mu, x) &= c_1 \nabla_{xx}^2 V_1(x) + \sum_{k=1}^K c_2 (\nabla_{xx}^2 V_2^{(k)}(x, \bar{y}^{(k)}) + \mu \nabla_{xx\bar{y}^{(k)}}^3 V_2^{(k)}(x, \bar{y}^{(k)}) \cdot (\bar{y}^{(k)})') \\
&= c_1 \nabla_{xx}^2 V_1(x) + \sum_{k=1}^K c_2 \nabla_{xx}^2 V_2^{(k)}(x, \bar{y}^{(k)}) \\
&= \frac{1}{\mu} (\nabla_{xx}^2 \eta(\mu, x) - C) \\
&= \preceq \frac{1}{\mu} \nabla_{xx}^2 \eta(\mu, x).
\end{aligned}$$

The conclusion follows since $\nabla_{xx}^2 \eta(\mu, x)$ is positive semidefinite. \square

Now we show the validity of the first inequality in Definition 2 (iii).

Lemma 3. *For any $\mu > 0$ and $x \in \mathcal{F}_0$, we have*

$$|\nabla_x \eta'(\mu, x)^\top [\xi]| \leq \sqrt{\frac{(n_1 c_1 + n_2 c_2)(1 + K)}{\mu}} \nabla_{xx}^2 \eta(\mu, x) [\xi, \xi], \quad \forall \xi \in \mathbb{R}^{n_1}.$$

Proof. This follows [5, Proof of Lemma 4]. \square

With Lemma 2 and Lemma 3 established, we have that Theorem 5 is true.

4 A Class of Volumetric Barrier Algorithms for Solving SQPs

In §3 we have established that the parametric functions $\eta(\mu, \cdot)$ is a strongly self-concordant family. In this section we introduce a class of volumetric barrier algorithms for solving (6, 7). This class, indexed by a parameter $\gamma \in (0, 1)$, is stated formally in Algorithm 1.

Our algorithm is initialized with a starting point $x^0 \in \mathcal{F}_0$ and a starting value $\mu^0 > 0$ for the barrier parameter μ . We use δ as a measure of the proximity of the current point x to the central path, and β as a threshold for that measure. If the current x is too far away from the central path in the sense that $\delta > \beta$, we apply Newton's method to find a point close to the central path. Then we reduce the value of μ by a factor γ and repeat the whole process until the value of μ is within the tolerance ϵ .

5 Complexity Analysis

Depending on the manner in which γ in Algorithm 1 is chosen, we have two classes of algorithms: short-step algorithms and long-step algorithms. In the next two subsections we present the complexity analysis for these two classes of algorithms.

5.1 Complexity of Short Step Algorithms

The i^{th} iteration of the short-step algorithms is performed as follows: at the beginning of the iteration, we have $\mu^{(i-1)}$ and $x^{(i-1)}$ on hand and $x^{(i-1)}$ is close to the center path, i.e. $\delta(\mu^{(i-1)}, x^{(i-1)}) \leq \beta$. After we reduce the parameter μ from $\mu^{(i-1)}$ to $\mu^i := \gamma \mu^{(i-1)}$, we have that $\delta(\mu^i, x^{(i-1)}) \leq 2\beta$. Then a full Newton step is taken to find a new point x^i with $\delta(\mu^i, x^i) \leq \beta$. We assume that we can solve all the subproblems exactly and we fix the value of $\gamma := 1 - 0.1/\sqrt{(n_1 c_1 + n_2 c_2)(1 + K)}$. We have the following complexity result:

Algorithm 1 A Class of Volumetric Barrier Algorithms for Solving SQP (6, 7)

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

$x := x^0$, $\mu := \mu^0$

while $\mu \geq \epsilon$ **do**

for $k = 1, 2, \dots, K$ **do**

 solve (12) to obtain $\bar{y}^{(k)}$

end for

 compute $\Delta x := -[\nabla_{xx}^2 \eta(\mu, x)]^{-1} \nabla_x \eta(\mu, x)$ using (18)

 compute $\delta(\mu, x) := \sqrt{\frac{1}{\mu} \Delta x^\top \nabla_{xx}^2 \eta(\mu, x) \Delta x}$ using (18)

while $\delta > \beta$ **do**

$x := x + \theta \Delta x$

for $k = 1, 2, \dots, K$ **do**

 solve (12) to obtain $\bar{y}^{(k)}$

end for

 compute $\Delta x := -[\nabla_{xx}^2 \eta(\mu, x)]^{-1} \nabla_x \eta(\mu, x)$ using (18)

 compute $\delta(\mu, x) := \sqrt{\frac{1}{\mu} \Delta x^\top \nabla_{xx}^2 \eta(\mu, x) \Delta x}$ using (18)

end while

$\mu := \gamma \mu$

end while

Theorem 6. Let $\beta := (2 - \sqrt{3})/2$ and $\gamma := 1 - 0.1/\sqrt{(n_1 c_1 + n_2 c_2)(1 + K)}$ in Algorithm 1. If $\delta(\mu^0, x^0) \leq \beta$, then short-step algorithms terminate with at most $O(\sqrt{(n_1 c_1 + n_2 c_2)(1 + K)} \ln(\mu^0/\epsilon))$ iterations.

Proof. This follows [5, Proofs of Lemma 5, 6] since we have shown that $\eta(\mu, \cdot)$ is a self-concordant family in Theorem 5. \square

5.2 Complexity of the Long Step Algorithms

In the long-step version of the algorithm, the factor $\gamma \in (0, 1)$ is arbitrarily chosen. It has potential for larger decrease on the objective function value, however, several damped Newton steps might be needed for recentering.

Suppose at the beginning of the i^{th} iteration of the algorithm we have a point $x^{(i-1)}$, which is sufficiently close to $x(\mu^{(i-1)})$, where $\mu^{(i-1)}$ is the current value for the barrier parameter μ and $x(\mu^{(i-1)})$ is the solution to (11) for $\mu := \mu^{(i-1)}$. We reduce the barrier parameter from $\mu^{(i-1)}$ to $\mu^i := \gamma \mu^{(i-1)}$, where $\gamma \in (0, 1)$, and we search for a point x^i that is sufficiently close to $x(\mu^i)$. We have the following complexity result:

Theorem 7. Let $\beta := 1/6$ and $\gamma \in (0, 1)$ be arbitrary in Algorithm 1. If $\delta(\mu^0, x^0) \leq \beta$, then long-step algorithms terminate with at most $O((n_1 c_1 + n_2 c_2)(1 + K) \ln(\mu^0/\epsilon))$ iterations.

Proof. This follows [5, Proofs of Lemma 7, 8, 9] since we have shown that $\eta(\mu, \cdot)$ is a self-concordant family in Theorem 5. \square

6 Concluding Remarks

In this paper we have derived a class of volumetric barrier decomposition algorithms for SQPs and shown that certain members of the class of algorithms have polynomial complexity. Our derivation and proofs made use of results obtained in [5] for SSDPs. It seems quite possible to extend the work we report in this paper to derive volumetric barrier decomposition algorithms for general stochastic convex programs. A forthcoming paper will report details of derivation of such a class of algorithms.

References

- [1] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.*, 5:13–51, 1995.
- [2] K. M. Anstreicher. Volumetric path following algorithms for linear programming. *Mathematical Programming*, 76(1), Ser. B, 245–263, 1997.
- [3] K. M. Anstreicher. The volumetric barrier for semidefinite programming. *Mathematics of Operations Research*, 25(3):365–380, 2000.
- [4] K. A. Ariyawansa and Y. Zhu. Stochastic Semidefinite Programming: A New Paradigm for Stochastic Optimization. *4OR—The Quarterly Journal of the Belgian, French and Italian OR Societies* (in press). Available as Technical Report 2004-10, Department of Mathematics, Washington State University, Pullman, WA, October 2004, at <http://www.math.wsu.edu/math/TRS/2004-10.pdf>.
- [5] K. A. Ariyawansa and Y. Zhu. *A Class of Polynomial Volumetric Barrier Decomposition Algorithms for Stochastic Semidefinite Programming* (submitted). Available as Technical Report 2006-7, Department of Mathematics, Washington State University, Pullman, WA, 2006, at <http://www.math.wsu.edu/math/TRS/2006-7.pdf>.
- [6] G. M. Cho. Log-barrier method for two-stage quadratic stochastic programming. *Appl. Math. Comput.*, 164, no. 1, 45–69, 2005.
- [7] A. J. Felt. *A Computational Evaluation of Interior Point Cutting Plane Algorithms for Stochastic Programs*. Ph.D. dissertation, Department of Pure and Applied Mathematics, Washington State University, Pullman, WA, August 2000.
- [8] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley, New York, NY, 1968.
- [9] P. L. Jiang. *Polynomial Cutting Plane Algorithms for Stochastic Programming and Related Problems*. Ph.D. dissertation, Department of Pure and Applied Mathematics, Washington State University, Pullman, WA, 1997.
- [10] O. L. Mangasarian. *Nonlinear Programming*. McGraw-Hill, New York, 1969.
- [11] Yu. E. Nesterov and A. S. Nemirovskii. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Publications. SIAM, Philadelphia, PA, 1994.
- [12] R. T. Rockafellar and R. J.-B Wets. A Lagrangian finite generation technique for solving linear-quadratic problems in stochastic programming. *Math. Programming Studies* 28:63–93,1996.

- [13] P. M. Vaidya. A new algorithm for minimizing convex functions over a convex set. *Math. Program., Ser. A*, 73:291–341, 1996.
- [14] G. Zhao. A log-barrier method with Benders decomposition for solving two-stage stochastic linear programs. *Math. Program., Ser. A*, 90:507–536, 2001.
- [15] Y. Zhu. *Semidefinite Programming Under Uncertainty*. Ph.D. dissertation, Department of Mathematics, Washington State University, Pullman, WA, August 2006.