SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

A method for weighted projections to the positive definite cone

T. Valkonen

SFB-Report No. 2012-016 August 2012

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

Der Wissenschaftsfonds.

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

A method for weighted projections to the positive definite cone

Tuomo Valkonen*

Abstract

We study the numerical solution of the problem $\min_{\mathbf{X}\geq\mathbf{0}} \|\mathbf{B}\mathbf{X}-\mathbf{c}\|_2$, where $\mathbf{X} \in \mathcal{S}_m$ is a symmetric square matrix, and $\mathbf{B} : \mathcal{S}_m \to \mathbb{R}^N$ is a linear operator, such that $\mathbf{B}^*\mathbf{B}$ is invertible. With ρ the desired fractional duality gap, we prove $O(\sqrt{m}\log\rho^{-1})$ iteration complexity for a simple primal-dual interior point method directly based on those for linear programs with semi-definite constraints, however not demanding the numerically expensive scalings inherent in these methods to force fast convergence.

Mathematics subject classification: 90C22, 90C51, 92C55.

Keywords: semi-definite, interior point, projection, quadratic programming diffusion tensor imaging.

1. Introduction

Let $\mathbf{B} : \mathcal{S}_m \to \mathbb{R}^N$ be a linear operator on symmetric $m \times m$ real matrices. Assume that $\mathbf{M} := \mathbf{B}^*\mathbf{B}$ is invertible. Given $\mathbf{c} \in \mathbb{R}^N$, we study the solution of the problem

$$\min_{\mathbf{X}\in\mathcal{S}_m^+} \|\mathbf{B}\mathbf{X} - \mathbf{c}\|_2, \tag{P}$$

with $\mathcal{S}_m^+ := \{\mathbf{X} \in \mathcal{S}_m \mid \mathbf{X} \geq \mathbf{0}\}$ denoting the set of positive semi-definite symmetric $m \times m$ matrices. Setting $\mathbf{C} := \mathbf{B}^*\mathbf{c}$, this is an instance of the quadratic optimisation problem with positive semi-definite constraints (quadratic SDP)

$$\min_{\mathbf{X}\in\mathcal{S}_m^+} \frac{1}{2}\langle\mathbf{X}, \mathbf{M}\mathbf{X}\rangle - \langle\mathbf{C}, \mathbf{X}\rangle, \tag{Q}$$

where we assume that the linear operator $\mathbf{M} : \mathcal{S}_m \to \mathcal{S}_m$ is self-adjoint and positive definite. More generally, both (P) and (Q) are instances of *semi-definite linear complementarity problems* (SDLCPs), discussed in, e.g., [9, 8].

To motivate the study of the instance (P), we first observe that if $\mathbf{B}\mathbf{X} = \mathrm{vec}(\mathbf{X})$ is the vectorisation of the matrix $\mathbf{X}$, then $\|\mathbf{B}\mathbf{X} - \mathbf{c}\|_2 = \|\mathbf{X} - \mathbf{C}\|_F$, where $\mathbf{C} := \mathrm{vec}^{-1}(\mathbf{c})$. Thus the solution $\hat{\mathbf{X}}$ of (P) is simply the projection of $\mathbf{C} \in \mathcal{S}_m$ to the positive semi-definite cone per the Frobenius norm. As a second source of motivation, we introduce an application from diffusion tensor imaging (DTI). [2, 6, 21] A diffusion tensor field $\mathbf{u} \in L^1(\Omega; \mathcal{S}_3)$ on a domain $\Omega \subset \mathbb{R}^3$ is determined by the Stejskal-Tanner equation

$$a_i(x) = a_0(x)\exp(-\langle\mathbf{b}_i \otimes \mathbf{b}_i, \mathbf{u}(x)\rangle), \quad (x \in \Omega), \tag{1.1}$$

from multiple diffusion weighted MRI (magnetic resonance imaging) measurements $a_i \in L^1(\Omega)$, for varying diffusion gradients $\mathbf{b}_i \in \mathbb{R}^3$, $(i = 1, \ldots, N)$, as well as the zero gradient $\mathbf{b}_0 = 0$. When there are at least six measurements, such that the matrices $\mathbf{B}_i := \mathbf{b}_i \otimes \mathbf{b}_i$ form a (necessarily non-orthogonal!)

* Institute for Mathematics and Scientific Computing, University of Graz, Austria. tuomo.valkonen@iki.fi

basis of $\mathcal{S}_3$, then $\mathbf{u}(x)$ can be determined by linear least-squares from (1.1). Namely, at each $x \in \Omega$, we solve

$$\min_{\mathbf{X}\in\mathcal{S}_3} \sum_{i=1}^{N} (\langle \mathbf{B}_i, \mathbf{X}\rangle - c_i)^2,$$

where $c_i := -\log(a_i(x)/a_0(x))$. Writing $\mathbf{B}\mathbf{X} := (\langle \mathbf{B}_1, \mathbf{X}\rangle, \ldots, \langle \mathbf{B}_N, \mathbf{X}\rangle)$ and $\mathbf{c} := (c_1, \ldots, c_N)$, this problem has the form

$$\min_{\mathbf{X}\in\mathcal{S}_m} \|\mathbf{B}\mathbf{X} - \mathbf{c}\|_2. \tag{1.2}$$

The solution is, of course, $\mathbf{X} = \mathbf{M}^{-1}\mathbf{B}^*\mathbf{c}$, provided $\mathbf{M}$ is invertible. This is how the diffusion tensor $\mathbf{u}(x) = \mathbf{X}$ is conventionally solved. [2] The only difference between (1.2) and (P) is that in the former there is no semi-definiteness constraint. But, *non-positive-definite diffusion-tensors are non-physical*, and the solution of (1.2) might not be positive definite, or even positive semi-definite. This is why it should be better, from an application point of view, to solve the problem (P) instead of (1.2). Our further goal is to incorporate a fidelity function based on (P) directly into the denoising framework developed in [21], instead of employing (1.2) as a first step. To facilitate massively parallel GPU (graphics processing unit) implementation, we look for a *simple* yet fast algorithm to solve (Q), hence (P), for small m.

Various methods exists for the solution of (Q), and more generally SDLCPs [9, 8, 4, 18, 11, 14, 22]. Many of these methods are rather involved predictor-corrector methods, while others have expensive steps. One example of a rather simple primal-dual path-following method is presented in [9] for general monotone SDLCPs, closely related to methods for linear programming on positive definite cones (linear SDP). [10, 12, 15, 13] This method employs the standard matrix product in the relaxed *complementarity condition* $\mathbf{X}\mathbf{S} = \sigma\mu\mathbf{I}$ between the primal variable $\mathbf{X}$ and the dual variable $\mathbf{S}$, and depends on a skew-symmetric modification of the search direction to ensure existence. Convergence is proved only in small neighbourhoods ($\gamma < 0.1$) of the central path, with complexity $O(\sqrt{m}\log\rho^{-1})$, for ρ the desired fractional duality gap.

In this paper, employing the symmetric product $\mathbf{X}\circ\mathbf{S} := (\mathbf{X}\mathbf{S} + \mathbf{S}\mathbf{X})/2$ for the relaxed complementarity condition $\mathbf{X}\circ\mathbf{S} = \sigma\mu I$, we derive another simple method for (P), (Q). It is the direct analogue of the methods [10, 17, 16, 1, 5] for linear SDPs and, more generally, linear programs on general symmetric cones derived through the Jordan-algebraic [3, 7] approach. Unlike in the linear SDP case, we will however see that we do not need to perform numerically expensive XS, SX, or Nesterov-Todd scalings of the primal and dual variables before linearization of the optimality conditions to ensure $O(\sqrt{m}\log\rho^{-1})$ iteration complexity. The precise iteration complexity that we derive will, however, depend on the spread of the eigenvalues of $\mathbf{M}$, namely the ratio Θ/θ, where $\Theta\mathbf{I} \geq \mathbf{M} \geq \theta\mathbf{I}$. The studied method and proofs could also be extended to other symmetric cones. For the sake of conciseness, we however concentrate on the positive definite cone.

In the rest of this short paper, we first study optimality conditions and search directions for (Q). We then introduce and study neighbourhoods of the central path in Section 3. We then show convergence of the proposed method in Section 4. Finally, in Section 5, we briefly present numerical results. Larger-scale numerical application will follow in an application-oriented follow-up work to [21].

2. Optimality conditions and search directions

We begin with a few definitions and known facts. We denote the symmetric (Jordan algebra) matrix product by

$$\mathbf{X}\circ\mathbf{S} := \frac{1}{2}(\mathbf{X}\mathbf{S} + \mathbf{S}\mathbf{X}).$$

For later use, we also define for each $X \in \mathcal{S}_m$ the corresponding linear operator $L(\mathbf{X}) : \mathcal{S}_m \to \mathcal{S}_m$ by

$$L(\mathbf{X})\mathbf{S} := \mathbf{X} \circ \mathbf{S}.$$

We know that $L(\mathbf{X})$ is self-adjoint, $\langle L(\mathbf{X})\mathbf{V}, \mathbf{W}\rangle = \langle L(\mathbf{X})\mathbf{W}, \mathbf{V}\rangle$ for $\mathbf{V}, \mathbf{W} \in \mathcal{S}_m$. It is moreover positive definite for $\mathbf{X} \in \operatorname{int} \mathcal{S}_m^+$, where the interior

$$\operatorname{int} \mathcal{S}_m^+ = \{\mathbf{X} \in \mathcal{S}_m \mid \mathbf{X} > \mathbf{0}\}$$

is the set of positive definite elements in $\mathcal{S}_m$. Hence $L(\mathbf{X})$ is, in particular, invertible for $\mathbf{X} \in \operatorname{int} \mathcal{S}_m^+$. [17, 3, 7]

For (Q) we now easily derive the optimality conditions

$$\mathbf{M}\mathbf{X} - \mathbf{S} = \mathbf{C}, \ \mathbf{X} \circ \mathbf{S} = \mathbf{0} \quad \text{for} \quad \mathbf{X}, \mathbf{S} \in \mathcal{S}_m^+, \tag{C}$$

where we call S the dual variable. To derive a Newton method for the solution of (Q), we first replace the *complementarity condition* $\mathbf{X} \circ \mathbf{S} = \mathbf{0}$ in (C) by the perturbed condition $\mathbf{X} \circ \mathbf{S} = \sigma\mu\mathbf{I}$, where $\mathbf{I}$ denotes the identity matrix, and $\sigma, \mu > 0$ are yet to be determined. This corresponds to replacing the constraint $\mathbf{X} \geq \mathbf{0}$ by the barrier function $-\sigma\mu \log\det(\mathbf{X})$ in (P). Thus the perturbed optimality conditions become

$$\mathbf{M}\mathbf{X} - \mathbf{S} = \mathbf{C}, \ \mathbf{X} \circ \mathbf{S} = \sigma\mu\mathbf{I} \quad \text{for} \quad \mathbf{X}, \mathbf{S} \in \mathcal{S}_m^+. \tag{C$_\mu$}$$

Suppose then that we have $\mathbf{X}, \mathbf{S} \in \mathcal{S}_m^+$ satisfying $\mathbf{M}\mathbf{X} - \mathbf{S} = \mathbf{C}$. We linearise (C$_\mu$) at $(\mathbf{X}, \mathbf{S})$, yielding the system

$$\mathbf{M}\Delta\mathbf{X} - \Delta\mathbf{S} = 0, \quad \text{and} \tag{L1}$$

$$\mathbf{X} \circ \Delta\mathbf{S} + \mathbf{S} \circ \Delta\mathbf{X} = \sigma\mu\mathbf{I} - \mathbf{X} \circ \mathbf{S} \tag{L2}$$

for the unknowns $\Delta\mathbf{X}, \Delta\mathbf{S} \in \mathcal{S}_m$. For a yet-undetermined step-size $\alpha > 0$ we set

$$\mathbf{X}(\alpha) := \mathbf{X} + \alpha\Delta\mathbf{X}, \quad \text{and} \quad \mathbf{S}(\alpha) := \mathbf{S} + \alpha\Delta\mathbf{S}, \tag{2.1}$$

as well as

$$\mu(\alpha) := \mu(\mathbf{X}(\alpha) \circ \mathbf{S}(\alpha)), \tag{2.2}$$

where, with a slight abuse of notation, we define

$$\mu(\mathbf{V}) := \sum_j \lambda_j(\mathbf{V})/m, \quad (\mathbf{V} \in \mathcal{S}_m),$$

for $\{\lambda_i(\mathbf{V})\}_{i=1}^m$ the eigenvalues of $\mathbf{V}$. Striving for a decrease $\mu(\alpha) < \mu(0)$ we pick $\sigma \in (0, 1)$ and $\mu := \mu(0) = \mu(\mathbf{X} \circ \mathbf{S})$ in (L2).

The questions now are the following: When is the system (L1), (L2) non-singular? What step-sizes α are valid to maintain in $\mathbf{X}(\alpha), \mathbf{S}(\alpha) \in \mathcal{S}_m^+$? How small can we make $\mu(\alpha)$? These bounds depend on $(\mathbf{X}, \mathbf{S})$ lying in a suitable neighbourhood of the *central path*, which are introduced and studied in the next section. First we, however, make a few remarks on linear programs and scaling.

Remark 2.1 (Linear programs with semi-definite constraints)**.** We recall that the aforementioned linear SDPs are of the form

$$\min \ \langle \mathbf{D}, \mathbf{X}\rangle \quad \text{such that } \mathbf{X} \in \mathcal{S}_m^+, \ \mathbf{A}\mathbf{X} = \mathbf{C}. \tag{2.3}$$

The optimality conditions may be expressed

$$\mathbf{A}\mathbf{X} = \mathbf{C}, \ \mathbf{A}^*\mathbf{y} + \mathbf{S} = \mathbf{D}, \ \mathbf{X} \circ \mathbf{S} = \mathbf{0} \quad \text{for} \quad \mathbf{X}, \mathbf{S} \in \mathcal{S}_m^+, \ \mathbf{y} \in \mathbb{R}^K. \tag{2.4}$$

In (2.4), in contrast to (C), the primal variable $\mathbf{X}$ and the dual variable $\mathbf{S}$ are coupled *only* through the complementarity condition $\mathbf{X} \circ \mathbf{S} = \mathbf{0}$. Moreover, when (2.4) is linearised at $(\mathbf{X}, \mathbf{S}, \mathbf{y})$, to solve for the step $(\Delta\mathbf{X}, \Delta\mathbf{S}, \Delta y)$, the first two conditions become $\mathbf{A}\Delta\mathbf{X} = 0$, and $\mathbf{A}^*\Delta\mathbf{y} + \Delta\mathbf{S} = 0$. From this it follows that $\langle \Delta\mathbf{X}, \Delta\mathbf{S} \rangle = 0$. This also fails for (C), and is the crucial ingredient in the established convergence proofs for established primal-dual interior point methods for (2.3); see [17] among the other references above. Under mild conditions on the operator $\mathbf{M}$, we will, however, be able to show fast convergence of the iteration (2.1), (2.2), along similar lines as was taken in [20, 19] to study convergence properties for optimality conditions of the type

$$\mathbf{A}_1\mathbf{X} = \mathbf{C},\ \mathbf{A}_2^*\mathbf{y} + \mathbf{S} = \mathbf{D},\ \mathbf{X} \circ \mathbf{S} = \mathbf{0} \quad \text{for} \quad \mathbf{X}, \mathbf{S} \in \mathcal{S}_m^+,\ \mathbf{y} \in \mathbb{R}^K,$$

related to diff-convex programming.

Remark 2.2 (Scaling)**.** Choosing a scaling $0 < \mathbf{P} \in \mathcal{S}_m$, and defining the *quadratic presentation* $\mathbf{Q_P} : \mathcal{S}_m \to \mathcal{S}_m$ of $\mathbf{P}$ as

$$\mathbf{Q_P}(X) := \mathbf{PXP},$$

we may also write (C_μ) in terms of $\widetilde{\mathbf{X}} := \mathbf{Q_P}\mathbf{X}$, $\underset{\sim}{\mathbf{S}} := \mathbf{Q_P}^{-1}\mathbf{X}$, and $\underset{\sim}{\mathbf{M}} := \mathbf{Q_P}^{-1}\mathbf{M}\mathbf{Q_P}^{-1}$ as

$$\underset{\sim}{\mathbf{M}}\widetilde{\mathbf{X}} - \underset{\sim}{\mathbf{S}} = \underset{\sim}{\mathbf{C}},\ \widetilde{\mathbf{X}} \circ \underset{\sim}{\mathbf{S}} = \sigma\mu\mathbf{I} \quad \text{for} \quad \widetilde{\mathbf{X}}, \underset{\sim}{\mathbf{S}} \in \mathcal{S}_m^+. \tag{$\widetilde{\mathrm{C}}_\mu$}$$

In many primal-dual interior point methods for linear SDPs [10, 17], this type of scaling is typically performed at each step to force fast convergence. The idea is to choose $\mathbf{P}$ such that $\mathbf{X}$ and $\mathbf{S}$ operator-commute, $L(\mathbf{X})L(\mathbf{S}) = L(\mathbf{S})L(\mathbf{X})$. For the XS-method $\mathbf{P} = \mathbf{X}^{1/2}$, for the SX-method $\mathbf{P} = \mathbf{S}^{1/2}$, and for the Nesterov-Todd method $\mathbf{P} = (\mathbf{Q}_{\mathbf{X}^{1/2}}(\mathbf{Q}_{\mathbf{X}^{1/2}}\mathbf{S})^{-1/2})^{-1/2}$. [17] We will not need this type of computationally expensive (matrix square root!) scalings for fast convergence. However our bounds will depend on $\mathbf{M}$. If we did perform scaling, we could obtain bounds that do not depend on $\mathbf{M}$, following the proof of [17].

Remark 2.3 (An easy special case)**.** Suppose that $\mathbf{MX} = \mathbf{Q_A}(\mathbf{X}) = \mathbf{AXA}$ for some $\mathbf{A} \in \operatorname{int} \mathcal{S}_m^+$. Then $\mathbf{M}\mathcal{S}_m^+ = \mathcal{S}_m^+$. [17, 3, 7] In the previous remark, let us choose $\mathbf{P} = \mathbf{A}^{1/2}$. The optimality conditions $(\widetilde{\mathrm{C}}_\mu)$ for $\mu = 0$ then become

$$\widetilde{\mathbf{X}} - \underset{\sim}{\mathbf{S}} = \underset{\sim}{\mathbf{C}},\ \widetilde{\mathbf{X}} \circ \underset{\sim}{\mathbf{S}} = \mathbf{0} \quad \text{for} \quad \widetilde{\mathbf{X}}, \underset{\sim}{\mathbf{S}} \in \mathcal{S}_m^+. \tag{2.5}$$

The solution $\widetilde{\mathbf{X}}$ of (2.5) is simply the Frobenius-norm projection of $\underset{\sim}{\mathbf{C}}$ to $\mathcal{S}_m^+$, and can easily be solved by projection of the eigenvalues $(\lambda_1(\underset{\sim}{\mathbf{C}}), \ldots, \lambda_m(\underset{\sim}{\mathbf{C}}))$ to $[0, \infty)^m$. Thus the solution $\mathbf{X} = \mathbf{Q}_{\mathbf{A}^{-1/2}}(\widetilde{\mathbf{X}})$ of (C) can also be easily calculated in this special case.

3. Neighbourhoods of the central path

Let now $\mathbf{P}_\mathbf{I}^\perp\mathbf{V} := \mathbf{V} - \mathbf{I}\mu(\mathbf{V})$ denote the projection of $\mathbf{V} \in \mathcal{S}_m$ to the subspace orthogonal to the identity $\mathbf{I}$. The spectrum of $\mathbf{P}_\mathbf{I}^\perp\mathbf{V}$ is then $\{\lambda_i(\mathbf{V}) - \mu(\mathbf{V})\}_{i=1}^m$.

With this, we now let $\gamma \in (0, 1)$, and define

$$\mathcal{N}_\bullet^*(\gamma) := \{(\mathbf{X}, \mathbf{S}) \in \operatorname{int} \mathcal{S}_m^+ \times \operatorname{int} \mathcal{S}_m^+ \mid \|\mathbf{P}_\mathbf{I}^\perp(\mathbf{X} \circ \mathbf{S})\|_\bullet \le \gamma\mu(\mathbf{X} \circ \mathbf{S})\} \tag{3.1}$$

for $\bullet \in \{F, 2, -\infty\}$. These correspond to the short-step, semi-long-step, and long-step neighbourhoods of $\mathcal{S}_m^+ \times \mathcal{S}_m^+$, and are obtained, respectively, with the Frobenius norm

$$\|\mathbf{V}\|_F := \sqrt{\textstyle\sum_{i=1}^m \lambda_i(\mathbf{V})^2} = \sqrt{\langle \mathbf{V}, \mathbf{V} \rangle},$$

the operator 2-norm

$$\|\mathbf{V}\|_2 := \max_{i=1,\dots,m} |\lambda_i(\mathbf{V})| = \max_{\mathbf{x}\in\mathbb{R}^m} \|\mathbf{V}\mathbf{x}\|_2/\|\mathbf{x}\|_2,$$

and, abusing norm notation for the sake of convenience, the function

$$\|\mathbf{V}\|_{-\infty} := -\min_i \lambda_i(\mathbf{V}).$$

For $\mathbf{P}_{\mathbf{I}}^{\perp}\mathbf{V}$ these have expressions

$$\begin{aligned}\|\mathbf{P}_{\mathbf{I}}^{\perp}\mathbf{V}\|_F &= \sqrt{\textstyle\sum_i (\lambda_i(\mathbf{V}) - \mu(\mathbf{V}))^2},\\ \|\mathbf{P}_{\mathbf{I}}^{\perp}\mathbf{V}\|_2 &= \max_i |\lambda_i(\mathbf{V}) - \mu(\mathbf{V})|, \quad \text{and}\\ \|\mathbf{P}_{\mathbf{I}}^{\perp}\mathbf{V}\|_{-\infty} &= \mu(\mathbf{V}) - \min \lambda_i(\mathbf{V}).\end{aligned} \tag{3.2}$$

It easily follows that

$$\mathcal{N}_F^*(\gamma) \subset \mathcal{N}_2^*(\gamma) \subset \mathcal{N}_{-\infty}^*(\gamma).$$

The following proposition is one of the crucial ingredients for our convergence proof.

Proposition 3.1. *Suppose* $(\mathbf{X}, \mathbf{S}) \in \mathcal{N}_{-\infty}^*(\gamma)$. *Then*

$$2(1-\gamma)\mu(\mathbf{X}\circ\mathbf{S}) \le \Big(\min_{i=1,\dots,m} \lambda_i(\mathbf{X}+\mathbf{S})\Big)^2. \tag{3.3}$$

Proof. We get from (3.2), (3.1) that

$$(1-\gamma)\mu(\mathbf{X}\circ\mathbf{S}) \le \min_i \lambda_i(\mathbf{X}\circ\mathbf{S}). \tag{3.4}$$

This shows that $\mathbf{X}\circ\mathbf{S} \in \mathcal{S}_m^+$, wherefore also

$$\langle \mathbf{X}\mathbf{y}, \mathbf{S}\mathbf{y}\rangle = \langle \mathbf{y}, (\mathbf{X}\circ\mathbf{S})\mathbf{y}\rangle > 0, \quad (\mathbf{y} \in \mathbb{R}^m).$$

Therefore

$$\begin{aligned}\min_i \lambda_i(\mathbf{X}\circ\mathbf{S}) &= \min_{\|\mathbf{y}\|=1} \langle \mathbf{y}, (\mathbf{X}\circ\mathbf{S})\mathbf{y}\rangle\\ &\le \min_{\|\mathbf{y}\|=1} \frac{1}{2}\big(\|\mathbf{X}\mathbf{y}\|_2^2 + \|\mathbf{S}\mathbf{y}\|_2^2\big) + \langle \mathbf{X}\mathbf{y}, \mathbf{S}\mathbf{y}\rangle\\ &= \min_{\|\mathbf{y}\|=1} \frac{1}{2}\big(\|\mathbf{X}\mathbf{y} + \mathbf{S}\mathbf{y}\|_2^2\big)\\ &= \frac{1}{2}\Big(\min_i \lambda_i(\mathbf{X}+\mathbf{S})\Big)^2.\end{aligned}$$

In the final step we have used the fact that $\mathbf{X}+\mathbf{S} \ge \mathbf{0}$. Recalling (3.4), the claim (3.3) follows. □

4. The method and its convergence

We now begin to study rates of convergence for the proposed method, consisting of the updates (2.1), (2.2). We assume that the linear operator $\mathbf{M} : \mathcal{S}_m \to \mathcal{S}_m$ is self-adjoint and satisfies for some $0 < \theta \le \Theta < \infty$ the condition

$$\Theta\langle \mathbf{V}, \mathbf{V}\rangle \ge \langle \mathbf{V}, \mathbf{M}\mathbf{V}\rangle \ge \theta\langle \mathbf{V}, \mathbf{V}\rangle, \quad (\mathbf{V} \in \mathcal{S}_m). \tag{A-M}$$

We begin by computing bounds on α for staying within $\mathcal{N}_\bullet^*(\gamma)$. With the notation

$$L := \{(\mathbf{X}, \mathbf{S}) \in \mathcal{S}_m^+ \times \mathcal{S}_m^+ \mid \mathbf{M}\mathbf{X} - \mathbf{S} = \mathbf{C}\}$$

for the feasible set, we have the following lemma.

Lemma 4.1. *If* $(\mathbf{X},\mathbf{S}) \in \mathcal{N}^*_\bullet(\gamma) \cap L$ *for some* $\bullet \in \{F, 2, -\infty\}$*, then* $(\mathbf{X}(\alpha),\mathbf{S}(\alpha)) \in \mathcal{N}^*_\bullet(\gamma) \cap L$ *for* $\alpha \in [0,\bar{\alpha}]$*, where*

$$\bar{\alpha} := \min\{1, \sigma\gamma\mu/(2\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F)\}. \tag{4.1}$$

Proof. This proof follows the outline of the proof of [17, Lemma 32], slightly modifying the proof to accommodate for the fact that $\langle \Delta\mathbf{X}, \Delta\mathbf{S}\rangle = 0$ does not hold; see also [20, 19].

Clearly $(\mathbf{X},\mathbf{S}) \in L$ and (L1) imply $(\mathbf{X}(\alpha),\mathbf{S}(\alpha)) \in L$. For the remainder, it will suffice to prove that

$$\|\mathbf{P}_\mathbf{I}^\perp(\mathbf{X}(\alpha) \circ \mathbf{S}(\alpha))\|_\bullet < \gamma\mu(\alpha) \quad \text{for} \quad \alpha \in (0,\bar{\alpha}]. \tag{4.2}$$

Indeed, irrespective of the choice of $\bullet$, (4.2) then holds for $\bullet = -\infty$ as well. Consequently

$$(1-\gamma)\mu(\alpha) < \min_i \lambda_i(\mathbf{X}(\alpha) \circ \mathbf{S}(\alpha)) = \min_i \lambda_i(\mathbf{X}(\alpha)\mathbf{S}(\alpha)),$$

where the last equality holds because $\mathbf{SX}$ and $\mathbf{XS}$ have the same spectrum for any $\mathbf{X},\mathbf{S} \in \mathcal{S}_m$. But then, taking the power of m on both sides, we get

$$((1-\gamma)\mu(\alpha))^m < \det(\mathbf{X}(\alpha)\mathbf{S}(\alpha)) = \det(\mathbf{X}(\alpha))\det(\mathbf{S}(\alpha)).$$

Now, by the continuity of the involved quantities in α, this condition would be violated if at some point $\alpha \in (0,\bar{\alpha}]$ one of the determinants were zero, that is, either $\mathbf{X}(\alpha)$ or $\mathbf{S}(\alpha)$ reached the boundary $\partial\mathcal{S}_m^+$. Thus (4.2) implies that $\mathbf{X}(\alpha),\mathbf{S}(\alpha) \in \operatorname{int}\mathcal{S}_m^+$ for $\alpha \in (0,\bar{\alpha}]$. Consequently, still by (4.2), $(\mathbf{X}(\alpha),\mathbf{S}(\alpha)) \in \mathcal{N}^*_\bullet(\gamma)$ for every $\alpha \in [0,\bar{\alpha}]$, as claimed.

It remains to prove (4.2). With the notation $\mathbf{Z} := \Delta\mathbf{X} \circ \Delta\mathbf{S}$, we have

$$\begin{aligned} \mu(\alpha) &= \mu(\mathbf{X} \circ \mathbf{S}) + \alpha\mu(\mathbf{X} \circ \Delta\mathbf{S} + \mathbf{S} \circ \Delta\mathbf{X}) + \alpha^2\mu(\Delta\mathbf{X} \circ \Delta\mathbf{S}) \\ &= \mu + \alpha(\sigma - 1)\mu + \alpha^2\mu(\mathbf{Z}) \\ &= (1-\alpha)\mu + \alpha\sigma\mu + \alpha^2\mu(\mathbf{Z}). \end{aligned} \tag{4.3}$$

as well as

$$\begin{aligned} \mathbf{P}_\mathbf{I}^\perp(\mathbf{X}(\alpha) \circ \mathbf{S}(\alpha)) &= \mathbf{P}_\mathbf{I}^\perp(\mathbf{X} \circ \mathbf{S}) + \alpha\mathbf{P}_\mathbf{I}^\perp(\mathbf{X} \circ \Delta\mathbf{S} + \mathbf{S} \circ \Delta\mathbf{X}) + \alpha^2\mathbf{P}_\mathbf{I}^\perp\mathbf{Z} \\ &= \mathbf{P}_\mathbf{I}^\perp(\mathbf{X} \circ \mathbf{S}) + \alpha\mathbf{P}_\mathbf{I}^\perp(\sigma\mu\mathbf{I} - \mathbf{X} \circ \mathbf{S}) + \alpha^2\mathbf{P}_\mathbf{I}^\perp\mathbf{Z} \\ &= (1-\alpha)\mathbf{P}_\mathbf{I}^\perp(\mathbf{X} \circ \mathbf{S}) + \alpha^2\mathbf{P}_\mathbf{I}^\perp\mathbf{Z}. \end{aligned} \tag{4.4}$$

To approximate the norm $\|\mathbf{P}_\mathbf{I}^\perp(\mathbf{X}(\alpha) \circ \mathbf{S}(\alpha))\|_\bullet$, for $\bullet = F$ we can use the triangle inequality on (4.4), whereas for $\bullet = 2, -\infty$, we apply, respectively, the inequalities

$$\begin{aligned} \max\lambda_i(\mathbf{V}+\mathbf{W}) &\le \max\lambda_i(\mathbf{V}) + \|\mathbf{W}\|_F, \quad \text{and} \\ -\min\lambda_i(\mathbf{V}+\mathbf{W}) &\le -\min\lambda_i(\mathbf{V}) + \|\mathbf{W}\|_F, \quad (\mathbf{V},\mathbf{W} \in \mathcal{S}_m). \end{aligned}$$

Therefore, for all $\bullet \in \{F, 2, -\infty\}$, we have the approximation

$$\begin{aligned} \|\mathbf{P}_\mathbf{I}^\perp(\mathbf{X}(\alpha) \circ \mathbf{S}(\alpha))\|_\bullet &\le (1-\alpha)\|\mathbf{P}_\mathbf{I}^\perp(\mathbf{X} \circ \mathbf{S})\|_\bullet + \alpha^2\|\mathbf{P}_\mathbf{I}^\perp\mathbf{Z}\|_F \\ &\le (1-\alpha)\gamma\mu + \alpha^2\|\mathbf{P}_\mathbf{I}^\perp\mathbf{Z}\|_F. \end{aligned}$$

Comparing this approximation against the expansion (4.3) of $\mu(\alpha)$, we find that (4.2) holds if

$$\alpha^2\|\mathbf{P}_\mathbf{I}^\perp\mathbf{Z}\|_F < (1 - \alpha - |1-\alpha| + \alpha\sigma)\gamma\mu + \gamma\alpha^2\mu(\mathbf{Z}).$$

Minding that $\mu(\mathbf{Z}) > 0$ by (L1), and that $\|\mathbf{P}_\mathbf{I}^\perp\mathbf{Z}\|_F \le 2\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F$, this follows if

$$2\alpha^2\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F \le (1 - \alpha - |1-\alpha| + \alpha\sigma)\gamma\mu.$$

The latter clearly holds for $\alpha \in [0,\bar{\alpha}] \subset [0,1]$. This completes the proof of (4.2) and the lemma. ☐

The following lemma provides the estimate on the fractional decrease $\mu(\bar{\alpha})/\mu$.

Lemma 4.2. *Assume the conditions of Lemma 4.1, and suppose* $0 < \sigma < m/(m+\gamma)$. *Then*

$$\begin{aligned}\delta := 1 - \mu(\bar{\alpha})/\mu &\geq (1-\sigma)\bar{\alpha}/2 \\ &= (1-\sigma)/2 \cdot \min\{1, \sigma\gamma\mu/(2\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F)\}.\end{aligned} \tag{4.5}$$

Proof. Observe that, thanks to the condition $\sigma < m/(m+\gamma)$, we have

$$\begin{aligned}(1-\sigma)m\mu/(2\langle\Delta\mathbf{X},\Delta\mathbf{S}\rangle) &\geq \sigma\gamma\mu/(2\langle\Delta\mathbf{X},\Delta\mathbf{S}\rangle) \\ &\geq \sigma\gamma\mu/(2\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F) \geq \bar{\alpha}.\end{aligned}$$

Therefore, using the expansion (4.3), we may calculate

$$1 - \mu(\bar{\alpha})/\mu = (1-\sigma)\bar{\alpha} - \bar{\alpha}^2\langle\Delta\mathbf{X},\Delta\mathbf{S}\rangle/(m\mu) \geq (1-\sigma)\bar{\alpha}/2. \tag{4.6}$$

Inserting (4.1) into (4.6) gives (4.5). □

Given a lower bound $\hat{\delta} \leq \delta$, a standard argument (see, e.g., [13]) shows that $\hat{\delta}^{-1}\log\rho^{-1}$ steps are sufficient to ensure that $\mu \leq \rho\overline{\mu}$ for an initial $\overline{\mu} > 0$ and desired decrease factor $\rho \in (0,1)$. To obtain the lower bound $\hat{\delta}$, and hence fast decrease in μ, by the previous lemma it suffices to bound $\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F/\mu$ from above. The standard proofs (see, e.g., [17]) for the semi-definite programming problem (2.3) rely at this point on the commutativity of $L(\widetilde{\mathbf{X}})$ and $L(\underset{\sim}{\mathbf{S}})$, where $\widetilde{\mathbf{X}}$ and $\underset{\sim}{\mathbf{S}}$ are the scaled variables as in Remark 2.2. We do not do so, as we want to avoid the computationally expensive scalings (involving matrix square roots), and show that we can avoid them, as do many other methods for quadratic SDPs and SDLCPs, referenced in the introduction.

Proposition 4.1 below is our most crucial ingredient for the convergence proof, and the main divergence from the proofs in [17] for linear SDPs. We show bounds for the short-step neighbourhood $\mathcal{N}_F^*(\gamma)$ and the semi-long-step neighbourhood $\mathcal{N}_2^*(\gamma)$.

Proposition 4.1. *Suppose* $\mathbf{M}$ *satisfies* (A-M), *and* $(\mathbf{X},\mathbf{S}) \in \mathcal{N}_F^*(\gamma)$. *Then*

$$\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F/\mu \leq \frac{((1-\sigma)\sqrt{m}+\gamma)^2}{2(1-\gamma)}(\Theta/\theta)^2. \tag{4.7}$$

If $(\mathbf{X},\mathbf{S}) \in \mathcal{N}_2^*(\gamma)$, *then*

$$\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F/\mu \leq \frac{(1-\sigma+\gamma)^2}{2(1-\gamma)}m(\Theta/\theta)^2. \tag{4.8}$$

Moreover, in both cases, there exists a unique solution $(\Delta\mathbf{X},\Delta\mathbf{S})$ *to the system* (L1), (L2).

Proof. Let us set

$$\mathbf{A} := L(\mathbf{S}) + L(\mathbf{X})\mathbf{M}.$$

Assuming that $\mathbf{A}$ is invertible, from (L1), (L2), we have

$$\Delta\mathbf{X} = \mathbf{A}^{-1}(\sigma\mu\mathbf{I} - \mathbf{X}\circ\mathbf{S}), \quad \text{and} \quad \Delta\mathbf{S} = \mathbf{M}\Delta\mathbf{X}. \tag{4.9}$$

Hence

$$\frac{\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F}{\mu} \leq \Theta\frac{\|\Delta\mathbf{X}\|_F^2}{\mu} \leq \Theta\frac{|\mathbf{A}^{-1}|_{\max}^2\|\sigma\mu\mathbf{I} - \mathbf{X}\circ\mathbf{S}\|_F^2}{\mu}, \tag{4.10}$$

where we denote

$$|\mathbf{A}^{-1}|_{\max} := \max\{\|\mathbf{A}^{-1}\mathbf{V}\|_F \mid \mathbf{V} \in \mathcal{S}_m,\ \|\mathbf{V}\|_F = 1\} \leq 1/|\mathbf{A}|_{\min}$$

and

$$|\mathbf{A}|_{\min} := \min\{\|\mathbf{AV}\|_F \mid \mathbf{V} \in \mathcal{S}_m,\ \|\mathbf{V}\|_F = 1\}.$$

We want to bound $|\mathbf{A}|_{\min}$ from below. Picking $\mathbf{V} \in \mathcal{S}_m$ with $\|\mathbf{V}\|_F = 1$, we calculate

$$\begin{aligned} \|\mathbf{AV}\|_F &= \|L(\mathbf{S}+\Theta\mathbf{X})L(\mathbf{S}+\Theta\mathbf{X})^{-1}\mathbf{AV}\|_F \\ &\geq |L(\mathbf{S}+\Theta\mathbf{X})|_{\min}\|L(\mathbf{S}+\Theta\mathbf{X})^{-1}\mathbf{AV}\|_F, \end{aligned} \tag{4.11}$$

We know that (see, e.g., [17])

$$|L(\mathbf{S}+\Theta\mathbf{X})|_{\min} = \min_i \lambda_i(\mathbf{S}+\Theta\mathbf{X}).$$

Observing that $(\mathbf{X},\mathbf{S}) \in \mathcal{N}^*_{-\infty}(\gamma)$ implies $(\Theta\mathbf{X},\mathbf{S}) \in \mathcal{N}^*_{-\infty}(\gamma)$, and referring to Proposition 3.1, we thus find that

$$|L(\mathbf{S}+\Theta\mathbf{X})|_{\min} \geq \sqrt{2(1-\gamma)\Theta\mu} \tag{4.12}$$

To estimate $\|L(\mathbf{S}+\Theta\mathbf{X})^{-1}\mathbf{AV}\|_F$, we observe that

$$L(\mathbf{S}+\Theta\mathbf{X})^{-1}\mathbf{A} = \mathbf{I} + L(\mathbf{S}+\Theta\mathbf{X})^{-1}L(\mathbf{X})(\mathbf{M}-\Theta\mathbf{I}).$$

By Lemma 4.3 below, we have

$$|L(\mathbf{S}+\Theta\mathbf{X})^{-1}L(\mathbf{X})|_{\max} \leq 1/\Theta.$$

Hence, employing (A-M) to estimate

$$|\mathbf{M}-\Theta\mathbf{I}|_{\max} \leq \Theta - \theta,$$

we obtain the bound

$$\begin{aligned} \|L(\mathbf{S}+\Theta\mathbf{X})^{-1}\mathbf{AV}\|_F &\geq 1 - \|L(\mathbf{S}+\Theta\mathbf{X})^{-1}L(\mathbf{X})(\mathbf{M}-\Theta\mathbf{I})\mathbf{V}\|_F. \\ &\geq 1 - |L(\mathbf{S}+\Theta\mathbf{X})^{-1}L(\mathbf{X})|_{\max}|\mathbf{M}-\Theta\mathbf{I}|_{\max} \\ &\geq 1 - (1/\Theta)(\Theta-\theta) \\ &= \theta/\Theta. \end{aligned} \tag{4.13}$$

Since $\mathbf{V} \in J$, $\|\mathbf{V}\|_F = 1$ was arbitrary, applying (4.12) and (4.13) in (4.11) yields

$$|\mathbf{A}|_{\min} \geq \|\mathbf{AV}\|_F \geq \sqrt{2(1-\gamma)\theta^2\mu/\Theta} > 0. \tag{4.14}$$

As this bound shows that $\mathbf{A}$ is invertible, the system (L1), (L2) has the unique solution (4.9). Moreover, applying (4.14) in (4.10) yields

$$\frac{\|\Delta\mathbf{X}\|_F\|\Delta\mathbf{S}\|_F}{\mu} \leq \frac{\|\sigma\mu\mathbf{I}-\mathbf{X}\circ\mathbf{S}\|_F^2}{2(1-\gamma)\mu^2}(\Theta/\theta)^2. \tag{4.15}$$

In the case $(\mathbf{X},\mathbf{S}) \in \mathcal{N}^*_F(\gamma)$, using the fact that $\|\mathbf{X}\circ\mathbf{S}-\mu\mathbf{I}\|_F \leq \gamma\mu$, and the triangle inequality, we have

$$\begin{aligned} \|\sigma\mu\mathbf{I}-\mathbf{X}\circ\mathbf{S}\|_F &\leq \|\sigma\mu\mathbf{I}-\mu\mathbf{I}\|_F + \|\mathbf{X}\circ\mathbf{S}-\mu\mathbf{I}\|_F \\ &\leq ((1-\sigma)\sqrt{m}+\gamma)\mu. \end{aligned} \tag{4.16}$$

Hence, (4.7) follows from (4.15).

In the case $(\mathbf{X},\mathbf{S}) \in \mathcal{N}^*_2(\gamma)$, using $\|\mathbf{X}\circ\mathbf{S}-\mu\mathbf{I}\|_2 \leq \gamma\mu$, we have

$$\|\mu\mathbf{I}-\mathbf{X}\circ\mathbf{S}\|_F \leq \sqrt{m}\|\sigma\mu\mathbf{I}-\mathbf{X}\circ\mathbf{S}\|_2 \leq \gamma\mu\sqrt{m},$$

so that, similarly to (4.16), we obtain

$$\|\sigma\mu\mathbf{I}-\mathbf{X}\circ\mathbf{S}\|_F \leq (1-\sigma+\gamma)\mu\sqrt{m}.$$

Hence (4.15) shows (4.8), concluding the proof. □

We needed the following lemma to prove Proposition 4.1.

Lemma 4.3. *Suppose* $\mathbf{X}, \mathbf{S} \in \operatorname{int} \mathcal{S}_m^+$, *and let* $\Theta > 0$. *Then*

$$|L(\mathbf{S} + \Theta\mathbf{X})^{-1} L(\mathbf{X})|_{\max} < 1/\Theta.$$

Proof. We have to prove that

$$\|L(\mathbf{S} + \Theta\mathbf{X})^{-1} L(\mathbf{X})\mathbf{V}\|_F < \|\mathbf{V}\|_F/\Theta, \quad (\mathbf{V} \in \mathcal{S}_m).$$

Squaring, expanding, and reorganising, this is written

$$\Theta^2 \langle L(\mathbf{X}) L(\mathbf{S} + \Theta\mathbf{X})^{-2} L(\mathbf{X})\mathbf{V}, \mathbf{V}\rangle < \langle \mathbf{V}, \mathbf{V}\rangle, \quad (\mathbf{V} \in \mathcal{S}_m).$$

Writing $\mathbf{V} = L(\mathbf{X})^{-1} L(\mathbf{S} + \Theta\mathbf{X}) L(\mathbf{X})\mathbf{W}$, we get the equivalent condition

$$\Theta^2 \langle L(\mathbf{X})^2 \mathbf{W}, \mathbf{W}\rangle < \langle L(\mathbf{X}) L(\mathbf{S} + \Theta\mathbf{X}) L(\mathbf{X})^{-2} L(\mathbf{S} + \Theta\mathbf{X}) L(\mathbf{X})\mathbf{W}, \mathbf{W}\rangle \tag{4.17}$$

for all $\mathbf{W} \in \mathcal{S}_m$. Expanding, we have

$$\begin{aligned} L(\mathbf{X}) L(\mathbf{S} + \Theta\mathbf{X}) & L(\mathbf{X})^{-2} L(\mathbf{S} + \Theta\mathbf{X}) L(\mathbf{X}) \\ &= L(\mathbf{X}) L(\mathbf{S}) L(\mathbf{X})^{-2} L(\mathbf{S}) L(\mathbf{X}) \\ &\quad + \Theta(L(\mathbf{X}) L(\mathbf{S}) + L(\mathbf{S}) L(\mathbf{X})) \\ &\quad + \Theta^2 L(\mathbf{X})^2. \end{aligned} \tag{4.18}$$

We know from the proof of [5, Corollary 4.4] that

$$\langle \big(L(\mathbf{X}) L(\mathbf{S}) + L(\mathbf{S}) L(\mathbf{X})\big)\mathbf{W}, \mathbf{W}\rangle > \langle L(\mathbf{S} \circ \mathbf{X})\mathbf{W}, \mathbf{W}\rangle > 0$$

under the condition $\mathbf{X}, \mathbf{S} \in \operatorname{int} \mathcal{S}_m^+$. Moreover, clearly by symmetricity

$$\langle L(\mathbf{X}) L(\mathbf{S} + \Theta\mathbf{X}) L(\mathbf{X})^{-2} L(\mathbf{S} + \Theta\mathbf{X}) L(\mathbf{X})\mathbf{W}, \mathbf{W}\rangle > 0.$$

Therefore, the expansion (4.18) shows (4.17), concluding the proof. □

We now concentrate on convergence in the case $(\mathbf{X}, \mathbf{S}) \in \mathcal{N}_F^*(\gamma)$. Recalling from (4.5) that

$$\delta \geq (1 - \sigma)/2 \cdot \min\{1, \sigma\gamma\mu/(2\|\Delta\mathbf{X}\|_F \|\Delta\mathbf{S}\|_F)\},$$

provided that $0 < \sigma \leq m/(m + \gamma)$, we now find with (4.7) that

$$\delta \geq (1 - \sigma)/2 \cdot \min\left\{1, \frac{\sigma\gamma(1 - \gamma)}{((1 - \sigma)\sqrt{m} + \gamma)^2} (\theta/\Theta)^2\right\}. \tag{4.19}$$

If we pick $\sigma = 1 - \gamma/\sqrt{m}$, which can be seen to satisfy $1 - \gamma \leq \sigma \leq m/(m + \gamma)$, we obtain

$$\begin{aligned} \delta &\geq \gamma/(2\sqrt{m}) \cdot \min\left\{1, \frac{(1 - \gamma/\sqrt{m})(1 - \gamma)}{4\gamma} (\theta/\Theta)^2\right\} \\ &\geq m^{-1/2} \min\left\{\frac{\gamma}{2}, \frac{(1 - \gamma)^2}{8} (\theta/\Theta)^2\right\}. \end{aligned}$$

If $\gamma = 1/2$, we in particular get

$$\delta \geq m^{-1/2} (\theta/\Theta)^2/32.$$

The results of this section are summarised in the following algorithm description and theorem.

Algorithm 4.1. The proposed interior point method is as follows.

1. Choose target accuracy $\underline{\mu} > 0$, and parameters $\gamma \in (0, 1)$, $\sigma \in (0, m/(m + \gamma)]$. Pick an initial iterate $(\mathbf{X}^0, \mathbf{S}^0) \in \mathcal{N}_\bullet(\gamma) \cap L$ for some choice of neighbourhood $\bullet \in \{F, 2\}$. Calculate $\mu^0 = \langle \mathbf{X}^0, \mathbf{S}^0 \rangle / m$, and set $i := 0$.
2. Solve $\Delta\mathbf{X}^i$ from
$$\big(L(\mathbf{S}^i) + L(\mathbf{X}^i)\mathbf{M}\big)\Delta\mathbf{X}^i = \sigma\mu^i\mathbf{I} - \mathbf{X}^i \circ \mathbf{S}^i, \tag{4.20}$$
and set
$$\Delta\mathbf{S}^i = \mathbf{M}\Delta\mathbf{X}^i.$$
3. Calculate the step length
$$\bar{\alpha} := \min\{1, \sigma\gamma\mu/(2\|\Delta\mathbf{X}^i\|_F\|\Delta\mathbf{S}^i\|_F)\},$$
and update $\mathbf{X}^{i+1} := \mathbf{X}^i + \bar{\alpha}\Delta\mathbf{X}^i$, and $\mathbf{S}^{i+1} := \mathbf{S}^i + \bar{\alpha}\Delta\mathbf{S}^i$, as well as $\mu^{i+1} = \langle \mathbf{X}^{i+1}, \mathbf{S}^{i+1} \rangle / m$,
4. If $\mu^{i+1} < \underline{\mu}$, terminate. Otherwise continue from Step 2 with $i := i + 1$.

We now state the convergence proof for the short-step neighbourhood $\mathcal{N}_F(\gamma)$.

Theorem 4.1. *Suppose* (A-M) *holds. Let* $(\mathbf{X}^0, \mathbf{S}^0) \in \mathcal{N}_F(\gamma) \cap L$ *and* $\rho := \mu^0/\underline{\mu}$. *Choose* $\sigma = 1 - \gamma/\sqrt{m}$. *Then in Algorithm 4.1,* $\mu^i < \underline{\mu}$ *for* $i > \delta^{-1} \log \rho^{-1}$, *where*
$$\delta = m^{-1/2} \min\left\{\frac{\gamma}{2}, \frac{(1-\gamma)^2}{8}(\theta/\Theta)^2\right\}. \tag{4.21}$$
In particular, if $\gamma = 1/2$, *then* $\mu^i < \underline{\mu}$ *for* $i > 32\sqrt{m}(\Theta/\theta)^2 \log \rho^{-1}$.

Proof. Follows from the discussion above. □

Remark 4.1. For the semi-long-step neighbourhood $\mathcal{N}_2(\gamma)$ we get for any $\gamma \in (0, 1)$ and $\sigma \in (0, m/(m + \gamma))$ by application of (4.5), (4.8) the bound $i > \zeta m(\Theta/\theta)^2 \log \rho^{-1}$ for some constant $\zeta = \zeta(\gamma, \sigma)$.

Remark 4.2. In practice, we may perform initialisation as follows. We pick $\mathbf{X}^0 = \beta I$ for a yet unknown $\beta > 0$. Then $\mathbf{S}^0 = \beta\mathbf{M}\mathbf{I} - \mathbf{C}$. The condition $(\mathbf{X}^0, \mathbf{S}^0) \in \mathcal{N}_\bullet(\gamma)$, namely
$$\|\mathbf{P}_{\mathbf{I}}^{\perp}(\mathbf{X}^0 \circ \mathbf{S}^0)\|_F \leq \gamma\mu(\mathbf{X}^0 \circ \mathbf{S}^0)$$
gives after division by β the condition
$$\|\beta\mathbf{P}_{\mathbf{I}}^{\perp}(\mathbf{M}\mathbf{I}) - \mathbf{P}_{\mathbf{I}}^{\perp}\mathbf{C}\|_F \leq \beta\mu(\mathbf{M}\mathbf{I}) - \mu(\mathbf{C}).$$
Squaring both sides, we get a second-order polynomial equation on β, from which we get a lower bound β_1 on β. Another lower bound β_2 on β is given by the condition $\mathbf{S}^0 \geq 0$. We may then choose $\beta := \max\{\beta_1, \beta_2, 0\}$.

Remark 4.3. An alternative initialisation strategy is to solve $\bar{\mathbf{X}}^0$ from $\mathbf{M}\bar{\mathbf{X}}^0 = \mathbf{C}$. Then we check if already $\bar{\mathbf{X}}^0 \geq \mathbf{0}$ (by Sylvester's criterion), in which case we may skip Algorithm 4.1. This is crucial for efficiency in our intended DTI denoising application. If $\bar{\mathbf{X}}^0 \not\geq \mathbf{0}$, we set $\mathbf{X}^0 := \bar{\mathbf{X}}^0 + \beta\mathbf{I}$ for an unknown $\beta > 0$, and $\mathbf{S}^0 := \mathbf{M}\mathbf{X}^0 - \mathbf{C}$. Then we calculate lower bounds for β analogously to Remark 4.2.

5. Numerical results

We tested the actual performance of the proposed method on the problem (P) numerically. As a first test case we took the identity operator $\mathbf{M} = \mathbf{B} = \mathbf{I}$ (in which case we could just use the QR algorithm to perform projection! – recall Remark 2.3) for varying dimensions m with the parameters $\gamma = 0.5$ and $\sigma = 1 - \gamma/\sqrt{m}$. The results of these experiments, for a decrease $\rho = 0.001$ of μ are plotted in Figure 1a, along with the iteration complexity upper bound from Theorem 4.1. The reported iteration count is the median over 10 samples of $\mathbf{C} \in \mathcal{S}_m$, with each component drawn independently from the uniform distribution on $[-1, 1]$. The maximum and minimum in the test run have a difference of at most ± 1 to the median. As we see, the method performs better than the prediction. However, as m becomes large, solving the system (4.20) (by LU decomposition) at each step starts to become prohibitively expensive, so other approaches are needed, such as the inexact steps of [18].

We are, however, mostly interested in the case $m = 3$, for which we report the iteration counts for varying parameters γ and σ in Figure 1b. Defining the symmetric presentation $\mathbf{PA}$ of

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{1,2} & a_{2,2} & a_{2,3} \\ a_{1,3} & a_{2,3} & a_{3,3} \end{bmatrix} \in \mathcal{S}_3$$

by

$$\mathbf{PA} = (a_{1,1}, a_{1,2}, a_{2,2}, a_{1,3}, a_{2,3}, a_{3,3}),$$

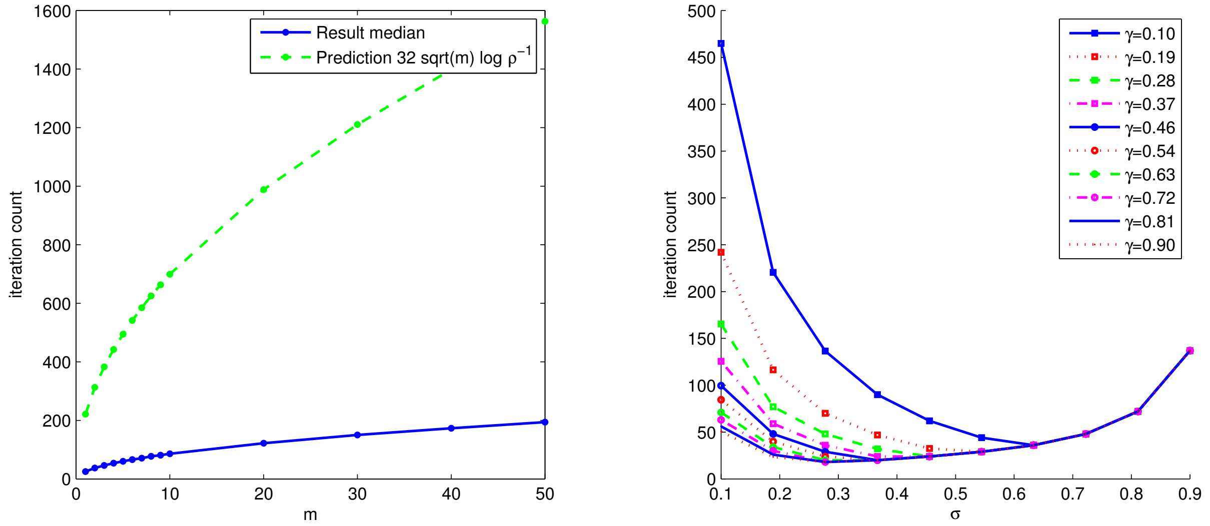
this time we use the fixed operator

$$\mathbf{M} = \mathbf{B}^*\mathbf{B} \approx \mathbf{P}^* \begin{bmatrix} 9.3716 & -0.0146 & 3.3252 & 0.0064 & -0.0062 & 3.2670 \\ -0.0146 & 3.3252 & 0.0107 & -0.0062 & -0.0028 & -0.0008 \\ 3.3252 & 0.0107 & 9.5023 & -0.0028 & 0.0565 & 3.2863 \\ 0.0064 & -0.0062 & -0.0028 & 3.2670 & -0.0008 & -0.0062 \\ -0.0062 & -0.0028 & 0.0565 & -0.0008 & 3.2863 & -0.0711 \\ 3.2670 & -0.0008 & 3.2863 & -0.0062 & -0.0711 & 9.3691 \end{bmatrix} \mathbf{P}$$

with $\Theta/\theta \approx 4.9$ and $\mathbf{B} : \mathcal{S}_3 \to \mathbb{R}^{52}$, constructed from a real DTI measurement setup. The reported data points are again the median over 10 samples of $\mathbf{c} \in \mathbb{R}^{52}$, with each component drawn independently uniformly from $[-1, 1]$. The difference of the maximum and minimum to the median is still too small to be observable in the figure. Apparently $\gamma = 0.9$ and $\sigma = 0.3$ would be a good choice of parameters, with a very small iteration count of about 20. This is significantly smaller than the theoretical bound of 147 given by (4.19). The choice $\sigma = 1 - \gamma/\sqrt{m} \approx 0.48$ also does not perform significantly worse. We may therefore conclude that the method performs very well for our intended purpose of "weighted projections to the positive definite cone" for small m.

References

1. F. Alizadeh and D. Goldfarb, *Second-order cone programming*, Mathematical Programming **95** (2003), 3–51, doi:10.1007/s10107-002-0339-5.
2. P. J. Basser and D. K. Jones, *Diffusion-tensor MRI: theory, experimental design and data analysis – a technical review.*, NMR in Biomedicine **15** (2002), 456–467, doi:10.1002/nbm.783.
3. J. Faraut and A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, 1994.
4. L. Faybusovich, *Euclidean jordan algebras and interior-point algorithms*, Positivity **1** (1997), 331–357, doi:10.1023/A:1009701824047.
5. ———, *Linear systems in Jordan algebras and primal-dual interior-point algorithms*, Journal of Computational and Applied Mathematics **86** (1997), 149–175, doi:10.1016/S0377-0427(97)00153-2.
6. P. Kingsley, *Introduction to diffusion tensor imaging mathematics: Parts I-III*, Concepts in Magnetic Resonance Part A **28** (2006), 101–179, doi:10.1002/cmr.a.20048. 10.1002/cmr.a.20049, 10.1002/cmr.a.20050.
7. M. Koecher, *The Minnesota notes on Jordan algebras and their applications*, *Lecture Notes in Mathematics*, volume 1710, Springer-Verlag, Berlin, 1999.



(a) Real and predicted iteration counts for varying m with $\mathbf{M} = \mathbf{I}$ with fixed.

(b) Iteration counts for varying γ and σ and fixed $m = 3$ and non-identity $\mathbf{M}$.

Figure 1: Iteration counts of numerical experiments for $\rho = 0.001$ fractional decrease of μ. Each data point is the median over 10 samples of $\mathbf{C}$.

8. M. Kojima, M. Shida and S. Shindoh, *Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs*, Mathematical Programming **80** (1998), 129–160, doi:10.1007/BF01581723.
9. M. Kojima, S. Shindoh and S. Hara, *Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices*, SIAM Journal on Optimization **7** (1997), 86–125, doi:10.1137/S1052623494269035.
10. R. Monteiro, *Polynomial Convergence of Primal-Dual Algorithms for Semidefinite Programming Based on the Monteiro and Zhang Family of Directions*, SIAM Journal on Optimization **8** (1998), 797–812, doi:10.1137/S1052623496308618.
11. R. Monteiro and S. Wright, *Local convergence of interior-point algorithms for degenerate monotone LCP*, Computational Optimization and Applications **3** (1994), 131–155, doi:10.1007/BF01300971.
12. Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, SIAM, 1994.
13. Y. E. Nesterov and M. J. Todd, *Self-scaled barriers and interior-point methods for convex programming*, Mathematics of Operations Research **22** (1997), 1–42, doi:10.1137/S1052623495290209.
14. J.-W. Nie and Y.-X. Yuan, *A predictor-corrector algorithm for QSDP combining Dikin-type and Newton centering steps*, Annals of Operations Research **103** (2001), 115–133, doi:10.1023/A:1012994820412.
15. F. A. Potra and S. J. Wright, *Interior-point methods*, Journal of Computational and Applied Mathematics **124** (2000), 281 – 302, doi:10.1016/S0377-0427(00)00433-7.
16. S. H. Schmieta and F. Alizadeh, *Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones*, Mathematics of Operations Research **26** (2001), 543–564, doi:10.1287/moor.26.3.543.10582.
17. ———, *Extension of primal-dual interior point algorithms to symmetric cones*, Mathematical Programming **96** (2003), 409–438, doi:10.1007/s10107-003-0380-z.
18. K.-C. Toh, *An inexact primal-dual path following algorithm for convex quadratic SDP*, Mathematical Programming **112** (2008), 221–254, doi:10.1007/s10107-006-0088-y.
19. T. Valkonen, *Diff-convex combinations of Euclidean distances: a search for optima*, number 99 in Jyväskylä Studies in Computing, University of Jyväskylä, 2008. Ph.D Thesis.
20. ———, *A primal-dual interior point method for diff-convex problems on symmetric cones*, Optimization (2011), doi:10.1080/02331934.2011.585465. Published online.
21. T. Valkonen and F. Knoll, *Total generalised variation in diffusion tensor imaging*, SFB-Report 2012-003, Karl-Franzens University of Graz (2012).
URL `http://math.uni-graz.at/mobis/publications/SFB-Report-2012-003.pdf`
22. G. Wang and Y. Bai, *Primal-dual interior-point algorithm for convex quadratic semi-definite optimization*,

Nonlinear Analysis: Theory, Methods, & Applications **71** (2009), 3389–3402, doi:10.1016/j.na.2009.01.241.