Inertial Proximal Block Coordinate Method for a Class of Nonsmooth and Nonconvex Sum-of-Ratios Optimization Problems

Radu Ioan Boţ†, Minh N. Dao‡ and Guoyin Li§

November 19, 2020

Abstract

In this paper, we consider a class of nonsmooth and nonconvex sum-of-ratios fractional optimization problems with block structure. This model class encompasses a broad spectrum of nonsmooth optimization problems such as the energy efficiency maximization problem and the sparse generalized eigenvalue problem. We first show that these problems can be equivalently reformulated as non-fractional nonsmooth optimization problems and propose an inertial proximal block coordinate method for solving them by exploiting the block structure of the underlying model. The global convergence of our method is guaranteed under the Kurdyka–Łojasiewicz (KL) property and some mild assumptions. We then identify the explicit exponents of the KL property for three important structured fractional optimization problems. In particular, for the sparse generalized eigenvalue problem with either cardinality regularization or sparsity constraint, we show that the KL exponents are $1/2$, and so, the proposed method exhibits linear convergence rate. Finally, we illustrate our theoretical results with both analytic and simulated numerical examples.

1. Introduction

We consider the following nonsmooth and nonconvex fractional optimization problem

$$\max_{x=(x_1, \ldots, x_m) \in S := S_1 \times \cdots \times S_m} F(x) := h(x_1, \ldots, x_m) + \sum_{i=1}^m f_i(x_i), \quad (P)$$

where, for each $i \in \{1, \ldots, m\}$, $\mathcal{H}_i$ is a finite-dimensional real Hilbert space, $S_i$ is a nonempty closed subset of $\mathcal{H}_i$, $h: \mathcal{H}_1 \times \cdots \times \mathcal{H}_m \to [-\infty, +\infty]$ is a (possibly) nonsmooth and nonconcave function, and $f_i, g_i: \mathcal{H}_i \to \mathbb{R}$ are locally Lipschitz functions such that, for all $x_i \in S_i$,

$$f_i(x_i) \geq 0 \quad \text{and} \quad g_i(x_i) > 0. \quad (1.1)$$

The model problem (P) covers various important optimization problems arising in diverse areas, such as the energy efficiency maximization problem and the sparse generalized eigenvalue problem. On the other hand, it belongs to the class of so-called sum-of-ratios optimization problems which are known as the most difficult problems in the fractional programming literature. Obviously, there is an alternative formulation for (P) which is obtained by replacing the maximum with minimum. Although these two formulations are in general independent (due to the non-negativity assumption (1.1)), the corresponding algorithmic development can be easily modified to suit the other formulation. Therefore, in this paper, we will focus on the maximum formulation. Below, we present a few explicit motivating examples for the model problem (P).

*Research of the first author is partially supported by the Austrian Science Fund (FWF), project number W1260-N35, and the research of the second and the third authors is partially supported by the Australian Research Council (ARC), project number DP190100555. Part of this research was done while the second author was with the Department of Applied Mathematics at the University of New South Wales.

†Faculty of Mathematics, University of Vienna, A-1090 Vienna, Austria. E-mail: radu.bot@univie.ac.at.

‡School of Engineering, Information Technology and Physical Sciences, Federation University Australia, Ballarat 3353, Australia. E-mail: m.dao@federation.edu.au.

§Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au.
Example 1.1 (Energy efficiency maximization problem [25]). Consider a power network which consists of \( m \) groups of links. Denote \( d_i \) the number of links in each group \( i \) of the network, and let \( d = \sum_{i=1}^{m} d_i \) be the total number of links in the network. For each group \( i \), the transmit power of a link \( j \) is denoted by \( x_{ij} \). Let \( x_i := (x_{i1}, \ldots, x_{id_i}) \), \( i \in \{1, \ldots, m\} \). We use \( \gamma_i(x_i) \) and \( P_i \) to denote the total signal to interference plus noise ratio (SINR) and the total power consumption experienced within the group \( i \), respectively. Accordingly, we can define the energy efficiency of the \( i \)-th group as \( f(\gamma_i(x_i)) / \mu_i x_i + P_i \), where \( f(\cdot) \) is a so-called efficiency function, \( \mu_i \geq 0 \) and \( P_i > 0 \). Then, the weighted energy efficiency (WEE) of the network is defined as
\[
\max_{x=(x_1, \ldots, x_m) \in \Delta_1 \times \cdots \times \Delta_m} \sum_{i=1}^{m} \frac{w_i f(\gamma_i(x_i))}{\mu_i x_i + P_i},
\]
where \( \Delta_i := \{x_i \in \mathbb{R}_{+}^{d_i} : x_{i}^{\text{min}} \leq x_{i} \leq x_{i}^{\text{max}} \} \) with \( 0 < x_{i}^{\text{min}} \leq x_{i}^{\text{max}} \). In particular, in the literature, the efficiency function is often taken as \( f(t) = \log(1 + t) \), while \( \gamma_i(x_i) = \mu_i x_i + r_i \) where \( u_i \in \mathbb{R}_{+}^{d_i} \setminus \{0\} \) and \( r_i \geq 0 \). In this case, the optimization problem becomes
\[
\max_{x=(x_1, \ldots, x_m) \in \Delta_1 \times \cdots \times \Delta_m} \sum_{i=1}^{m} \frac{w_i \log(1 + u_i^\top x_i + r_i)}{\mu_i x_i + P_i},
\]
and it obviously is a particular case of our model problem (P) for \( h \equiv 0 \), \( f_i(x_i) = w_i \log(1 + u_i^\top x_i + r_i) \), \( g_i(x_i) = \mu_i x_i + P_i \), and \( S_i = \Delta_i, \ i \in \{1, \ldots, m\} \).

Example 1.2 (Maximizing the sum of a quadratic function and the Rayleigh quotient over the unit sphere [26]). We consider the problem of maximizing the sum of a quadratic function and Rayleigh quotient over the unit sphere
\[
\max_{x \in \mathbb{R}^d} x^\top W x + \frac{x^\top A x}{x^\top B x} \quad \text{s.t.} \quad ||x|| = 1,
\]
where \( A, B \) are positive definite matrices and \( W \) is a symmetric matrix.

This problem arises in sparse Fisher discriminant analysis, in the context of which it is usually solved iteratively [26]. In particular, \( x \) is the desired discriminating vector in cluster analysis, the term \( x^\top W x \) is known as the Rayleigh quotient (or Fisher information in information science), and the quadratic term \( x^\top A x / x^\top B x \) serves as a local approximation for the sparse penalty term. Direct verification shows that this is a particular case of our model problem (P) for \( m = 1 \), \( h(x) = x^\top W x \), \( f_1(x) = x^\top A x \), \( g_1(x) = x^\top B x \) and \( S = \{x \in \mathbb{R}^d : ||x|| = 1\} \).

Example 1.3 (Sparse generalized eigenvalue problem). The generalized eigenvalue problem, which searches for the most dominant eigenvalues (or principal eigenvalues) and corresponding eigenvector, can be written as an optimization problem \( \max_{x \in \mathbb{R}^d} \{\lambda x^\top A x / x^\top B x : ||x|| = 1\} \). In numerical analysis, one seeks for an eigenvector with least number of nonzero entries, so that the information can be easily stored, explained and identified. This leads to a sparse generalized eigenvalue problem which can be formulated as
\[
\max_{x \in \mathbb{R}^d} \frac{x^\top A x}{x^\top B x} - \lambda \phi(x) \quad \text{s.t.} \quad ||x|| = 1.
\]
Here, \( A, B \) are symmetric matrices with \( A \) positive semidefinite and \( B \) positive definite, and \( \phi \) is a regularization function which induces sparsity of the solution. Typical choices of \( \phi \) include the \( \ell_0 \) regularization (or cardinality) function given by \( ||x||_0 = \{\text{number of } i : x_i \neq 0\} \), the \( \ell_1 \)-norm given by \( ||x||_1 = \sum_{i=1}^{d} |x_i| \), and the indicator function of the sparsity set \( C_r = \{x \in \mathbb{R}^d : ||x||_0 \leq r \} \) with \( r > 0 \). For example, in a recent study [24], the authors examined the sparse generalized eigenvalue problem with \( \phi(x) = \delta_{C_r}(x) \), where they proposed a truncated Rayleigh flow method (TRFM) and demonstrated the efficiency of this model problem on classification, correlation analysis and regression. Direct verification shows that the sparse generalized eigenvalue problem is a particular case of our model problem (P) with \( m = 1 \), \( h(x) = -\lambda \phi(x) \), \( f_1(x) = x^\top A x \), \( g_1(x) = x^\top B x \), and \( S = \{x \in \mathbb{R}^d : ||x|| = 1\} \).

In the case where \( m = 1 \) and \( h \equiv 0 \), problem (P) is known as the single ratio fractional programming problem \( \max_{x \in S} \frac{f_i(x)}{g_i(x)} \). A classical approach for solving the single ratio fractional programming problem is Dinkelbach’s
method and its iterative approximation version (see [9, 11]). In this approach, one typically constructs an iterative scheme which requires finding an optimal solution $x_{n+1}$ of the optimization problem

$$
\max_{x \in S} \{ f_1(x) - \theta_n g_1(x) \}
$$

(1.2)

in each iteration $n$, while $\theta_n$ is updated by $\theta_{n+1} := \frac{f_1(x_{n+1})}{g_1(x_{n+1})}$. For details of this approach, we refer the readers to [9, 11, 12, 22]. However, solving in each iteration an optimization problem of type (1.2) may be as expensive and difficult as solving the original problem in general. Recently, proximal type methods based on Dinkelbach’s approach have been proposed to tackle single ratio fractional programs [7, 8], where each subproblem is much easier to solve and sometimes has closed form solutions.

Unfortunately, in the case of sum-of-ratios fractional programs, that is either $m > 1$ or $h \neq 0$, Dinkelbach’s approach cannot be directly applied anymore. One naive approach is to convert the sum-of-ratios into single ratio’s cases and to apply Dinkelbach’s method. This approach increases the complexity of the function dramatically and leads to numerical methods with poor performance. For example, through this approach, a sum of three linear fractional functions becomes a fractional function whose numerator and denominator are degree 3 nonconvex polynomials, and so, the nice linearity structure is completely lost. Some important steps towards solving sum-of-ratios fractional programs are mainly limited to sum-of-ratios of linear or quadratic fractional programs, and rely on integer programming techniques such as branch and bound and convex relaxation methods, see for example [5, 17, 26]. These approaches, although highly appealing, are much less scalable than the proximal type methods, and cannot directly deal with settings in which nonsmooth functions are involved.

Despite this important progress, it is still no clear whether one can develop proximal methods for solving nonsmooth and nonconvex sum-of-ratios fractional programs (P) in the line of [7, 8] for single ratio cases. This forms the basic motivation of our work. Specifically, the contributions of this paper are as follows:

(1) In Section 3, we present an equivalent non-fractional program reformulation for the model problem (P) and establish the explicit relationship between global minimizers and stationary points of the two problems.

(2) In Section 4, we propose an inertial proximal method for solving the model problem (P). We then show that the iterative sequence generated by the proposed method is bounded and any of its limit points is a so-called lifted stationary point of problem (P). We also establish the convergence of the full sequence under the assumption that a suitable merit function satisfies the Kurdyka–Łojasiewicz (KL) property.

(3) In Section 5, we analyze several structured sum-of-ratios fractional programs and obtain the explicit KL exponents of the corresponding desingularization functions in the KL property: sum-of-ratios fractional quadratic programs with spherical constraint, generalized eigenvalue problems with cardinality regularization and generalized eigenvalue problems with sparsity constraints. In particular, we establish that, for the last two classes of fractional programs, the KL exponents are $1/2$. As a consequence, we obtain that the proposed numerical method exhibits linear convergence for these two classes of fractional programs.

(4) Finally, we illustrate the proposed method via both analytical and simulated numerical examples in Section 6.

2. Preliminaries

In this section, we recall some basic notations and preliminary results which will be used in this paper. We assume throughout that $H$, $H_1$, $\ldots$, $H_m$ are finite-dimensional real Hilbert spaces with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. The product space $H_1 \times \cdots \times H_m$ is also a real Hilbert space endowed with the inner product given by

$$
\langle (x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n) \rangle = \sum_{i=1}^{m} \langle x_i, y_i \rangle.
$$

The set of nonnegative integers is denoted by $\mathbb{N}$, the set of real numbers by $\mathbb{R}$, the set of nonnegative real numbers by $\mathbb{R}_+ := \{ x \in \mathbb{R} : x \geq 0 \}$, and the set of the positive real numbers by $\mathbb{R}_{++} := \{ x \in \mathbb{R} : x > 0 \}$.

The indicator function of a set $C$ is defined by $\delta_C(x) := 0$ if $x \in C$, and $\delta_C(x) := +\infty$ otherwise. Given an extended-real-valued function $f : H \to [-\infty, +\infty]$, its domain is defined by $\text{dom} f := \{ x \in H : f(x) < +\infty \}$. The
function \( f \) is proper if \( \text{dom} \ f \neq \emptyset \) and it never equals \(-\infty\). We say that \( f \) is lower semicontinuous if, for all \( x \in \mathcal{H} \), 
\[
 f(x) \leq \liminf_{y \to x} f(y),
\]
and convex if, for all \( x, y \in \text{dom} \ f \) and all \( \lambda \in (0, 1) \), 
\[
 f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y).
\]
The function \( f \) is said to be weakly convex (on \( \mathcal{H} \)) if there exists \( \alpha \geq 0 \) such that \( f + \frac{\alpha}{2} \| \cdot \|^2 \) is a convex function. The smallest constant \( \alpha \) such that \( f + \frac{\alpha}{2} \| \cdot \|^2 \) is convex is called the modulus of weak convexity. More generally, \( f \) is said to be weakly convex on \( S \subseteq \mathcal{H} \) with modulus \( \alpha \) if \( f + \delta_S \) is weakly convex with modulus \( \alpha \). Weakly convex functions form a broad class of functions which covers quadratic functions, convex functions and continuously differentiable functions whose gradient is Lipschitz continuous.

Let \( f: \mathcal{H} \to [-\infty, +\infty] \) and \( x \in \mathcal{H} \) with \( |f(x)| < +\infty \). The Fréchet subdifferential of \( f \) at \( x \) is given by 
\[
\partial f(x) := \left\{ u \in \mathcal{H} : \liminf_{z \to x} \frac{f(z) - f(x) - \langle u, z - x \rangle}{\| z - x \|} \geq 0 \right\},
\]
the limiting subdifferential of \( f \) at \( x \) is given by 
\[
\partial_L f(x) := \left\{ u \in \mathcal{H} : \exists x_n \xrightarrow{L} x, \ u_n \to u \ \text{with} \ u_n \in \partial f(x_n) \right\},
\]
and the horizon subdifferential of \( f \) at \( x \) is given by 
\[
\partial^\infty f(x) := \left\{ u \in \mathcal{H} : \exists x_n \xrightarrow{L} x, \ \lambda_n \downarrow 0, \ \lambda_n u_n \to u \ \text{with} \ u_n \in \partial f(x_n) \right\}.
\]
It follows from the above definition that the limiting subdifferential has the following robustness property 
\[
\partial_L f(x) = \left\{ u \in \mathcal{H} : \exists x_n \xrightarrow{L} x, \ u_n \to u \ \text{with} \ u_n \in \partial_L f(x_n) \right\}.
\]
The domain of \( \partial_L f \) is \( \text{dom} \partial_L f := \{ x \in \mathcal{H} : \partial_L f(x) \neq \emptyset \} \). If \( f \) is strictly differentiable\(^1\) at \( x \), then \( \partial_L f \) reduces to the derivative of \( f \), denoted by \( \nabla f \) (see [19, Corollary 1.82]). If \( f \) is convex, then the limiting subdifferential reduces to the classical subdifferential in convex analysis (see [19, Theorem 1.93]), i.e., 
\[
\partial_L f(x) = \left\{ u \in \mathcal{H} : \forall z \in \mathcal{H}, \ \langle u, z - x \rangle \leq f(z) - f(x) \right\}.
\]

In Appendix A, we collect some subdifferential rules and some calculations which will be of use in our analysis.

We end this section by recalling the celebrated Kurdyka–Łojasiewicz (KL) property [13, 15] which plays an important role in our convergence analysis later on. For each \( \eta \in (0, +\infty) \), let \( \Phi_\eta \) be the class of all continuous concave functions \( \varphi: [0, \eta] \to \mathbb{R}_+ \) such that \( \varphi(0) = 0 \) and \( \varphi \) is continuously differentiable with \( \varphi' > 0 \) on \((0, \eta)\).

A proper lower semicontinuous function \( f: \mathcal{H} \to (-\infty, +\infty] \) is said to satisfy the KL property [13, 15] at \( x \in \text{dom} \partial_L f \) if there exist a neighborhood \( U \) of \( x \), \( \eta \in (0, +\infty) \), and a function \( \varphi \in \Phi_\eta \) such that, for all \( x \in U \) with \( f(x) < f(x) < f(\mathcal{X}) + \eta \), one has 
\[
\varphi'(f(x) - f(x)) \text{dist}(0, \partial_L f(x)) \geq 1.
\]
If \( f \) satisfies the KL property at each point in \( \text{dom} \partial_L f \), then \( f \) is called a KL function. We say that \( f \) has the KL property at \( x \in \text{dom} \partial_L f \) with an exponent of \( \alpha \) if \( f \) satisfies the KL property at \( x \in \text{dom} \partial_L f \) and the corresponding function \( \varphi \) (often referred as desingularization function) can be chosen as \( \varphi(s) = \gamma s^{1-\alpha} \) for some \( \gamma \in \mathbb{R}_{++} \) and \( \alpha \in (0, 1) \). If \( f \) is a KL function and has the same exponent \( \alpha \) at any \( x \in \text{dom} \partial_L f \), then \( f \) is called a KL function with an exponent of \( \alpha \).

This definition encompasses broad classes of functions that arise in practical optimization problems. For example, it is known that if \( h \) is a proper lower semicontinuous semi-algebraic function, then \( f \) is a KL function with an exponent of \( \alpha \in (0, 1) \). The class of semi-algebraic functions covers many common nonsmooth functions that appear in modern optimization problems such as functions which can be written as a maximum or a minimum of finitely many polynomials, Euclidean norms, and the eigenvalues and the rank of a matrix. Also, sums, products, and quotients of semi-algebraic functions are semi-algebraic. As established in [1, 2, 6, 16] and in many subsequent works, the KL exponent is closely related to the rate of convergence of many commonly used optimization methods.

\(^1\)A function \( f \) is strictly differentiable at \( x \) if there exists \( u \in \mathcal{H} \) such that \( \lim_{y \to x} \frac{f(y) - f(x) - \langle u, y - x \rangle}{\| y - x \|} = 0 \). Clearly, if \( f \) is continuously differentiable at \( x \), then it is strictly differentiable at \( x \).
3. Equivalent non-fractional formulations

Below, we present a simple lemma which shows that a sum-of-ratios problem can be equivalently reformulated as a nonsmooth optimization problem which does not involve fractional functions. To do this, denote $\mathcal{H} = \mathcal{H}_1 \times \cdots \times \mathcal{H}_m$, $S = S_1 \times \cdots \times S_m$ and $y = (y_1, \ldots, y_m)$. We consider $$\max_{(x,y) \in S \times \mathbb{R}^m} h(x) + H(x, y) \quad \text{with} \quad H(x, y) := \sum_{i=1}^m \left[ 2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i) \right]. \quad (P_1)$$

We now show that problem $(P)$ and problem $(P_1)$ are equivalent. This property was mentioned in [5] for the case when $h \equiv 0$, and $f_i$ and $g_i$, $i \in \{1, \ldots, m\}$, are all affine-linear functions (that is, in the setting of fractional linear optimization). As its proof is elementary, we include it for the purpose of self-containment.

**Lemma 3.1 (Equivalent non-fractional optimization formulations).** Let $\mathbf{x} = (x_1, \ldots, x_m) \in \mathcal{H}_1 \times \cdots \times \mathcal{H}_m$. Then $\mathbf{x}$ is a global solution of $(P)$ if and only if $(\mathbf{x}, \mathbf{y})$ is a global solution for $(P_1)$, where $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$ and $y_i = \frac{\sqrt{f_i(x_i)}}{g_i(x_i)}$, $i \in \{1, \ldots, m\}$, in which case, both problems have the same optimal value.

**Proof.** Let $\mathbf{x}$ be a global solution of $(P)$. Then $\mathbf{x} \in S$ and, for all $\mathbf{x} \in S$,

$$h(x) + \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)} \leq h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})}.$$  

From the definition of $y_i$, $i \in \{1, \ldots, m\}$, one has

$$H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m \left( 2y_i \sqrt{f_i(\mathbf{x})} - y_i^2 g_i(x_i) \right) = \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)}. \quad (3.1)$$

Note that for each fixed $x_i \in S_i$, $y_i \mapsto H_i(x_i, y_i) := 2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i)$ is a strongly concave one-variable quadratic function. Direct verification shows that

$$\arg\max_{y_i \in \mathbb{R}} H_i(x_i, y_i) = \arg\max_{y_i \in \mathbb{R}} \{ 2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i) \} = \left\{ \frac{\sqrt{f_i(x_i)}}{g_i(x_i)} \right\}.$$  

Then, for all $(\mathbf{x}, \mathbf{y}) \in S \times \mathbb{R}^m$,

$$H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m H_i(x_i, y_i) \leq \sum_{i=1}^m \max_{y_i \in \mathbb{R}} H_i(x_i, y_i) = \sum_{i=1}^m H_i \left( x_i, \frac{\sqrt{f_i(x_i)}}{g_i(x_i)} \right) = \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)},$$

which implies that

$$h(\mathbf{x}) + H(\mathbf{x}, \mathbf{y}) \leq h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)} \leq h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} = h(\mathbf{x}) + H(\mathbf{x}, \mathbf{y}).$$

So, $(\mathbf{x}, \mathbf{y})$ is a global solution for $(P_1)$.

Conversely, let $(\mathbf{x}, \mathbf{y})$ be a global solution for $(P_1)$. Then, for all $(\mathbf{x}, \mathbf{y}) \in S \times \mathbb{R}^m$,

$$h(\mathbf{x}) + H(\mathbf{x}, \mathbf{y}) \leq h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} = h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)}.$$  

This shows that, for all $\mathbf{x} \in S$,

$$h(\mathbf{x}) + \max_{\mathbf{y} \in \mathbb{R}^m} H(\mathbf{x}, \mathbf{y}) \leq h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})}.$$  

Notice that

$$\max_{\mathbf{y} \in \mathbb{R}^m} H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m \max_{y_i \in \mathbb{R}} \{ 2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i) \} = \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)}.$$  

So, $\mathbf{x}$ is a global solution for $(P)$.

Finally, the conclusion that $(P)$ and $(P_1)$ have the same optimal value follows from (3.1).
Definition 3.2 (Stationary points). We say that $\mathbf{x} = (x_1, \ldots, x_m) \in S$ is a stationary point for (P) if $0 \in \partial_L (-F + \delta_S)(\mathbf{x})$, and a lifted stationary point for (P) if

$$0 \in \partial_L (-h + \delta_S)(\mathbf{x}) + \left( \begin{array}{c} -g_1(x_1) \partial_L f_1(x_1) + f_1(x_1) \partial_L g_1(x_1) \\ \vdots \\ -g_m(x_m) \partial_L f_m(x_m) + f_m(x_m) \partial_L g_m(x_m) \end{array} \right).$$

We also say that $(\mathbf{x}, \mathbf{y}) \in S \times \mathbb{R}^m$ is a stationary point for $(P_1)$ if $0 \in \partial_L (-h + \delta_S)(\mathbf{x}) + \partial_L (H)(\mathbf{x}, \mathbf{y})$, or equivalently, if $0 \in \partial_L (-h + \delta_S)(\mathbf{x}) + \partial^L_2(H)(\mathbf{x}, \mathbf{y})$ and $\mathbf{y} = (\sqrt{f_1(x_1)}/g_1(x_1), \ldots, \sqrt{f_m(x_m)}/g_m(x_m))$. Here $\partial^L_2$ denotes the subdifferential with respect to the $\mathbf{x}$-variable.

The next lemma summarises the relationship between the notion of a stationary point for (P), a lifted stationary point for (P) and a stationary point for $(P_1)$.

Lemma 3.3. Let $H_i(x_i, y_i) := 2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i)$, $i \in \{1, \ldots, m\}$, $\mathbf{x} = (x_1, \ldots, x_m) \in \mathcal{H}$ and $\mathbf{y} = (y_1, \ldots, y_m) \in \mathbb{R}^m$ with $y_i = \sqrt{f_i(x_i)}$. Suppose that $-h$ is proper lower semicontinuous and finite at $\mathbf{x}$ and that, for each $i \in \{1, \ldots, m\}$, $f_i$ and $g_i$ are Lipschitz continuous around $x_i$, $f_i(x_i) \geq 0$ and $g_i(x_i) > 0$. Then the following statements hold:

(i) If $\mathbf{x}$ is a stationary point for (P) and, for each $i \in \{1, \ldots, m\}$, $\partial f_i$ is nonempty-valued around $x_i$, then it is a lifted stationary point for (P). Conversely, if $\mathbf{x}$ is a lifted stationary point for (P) and, for each $i \in \{1, \ldots, m\}$, $f_i$ and $g_i$ are strictly differentiable at $x_i$, then $\mathbf{x}$ is a stationary point for (P).

(ii) Suppose that, for each $i \in \{1, \ldots, m\}$, $f_i(x_i) > 0$. If $(\mathbf{x}, \mathbf{y})$ is a stationary point for $(P_1)$ and, for each $i \in \{1, \ldots, m\}$, $\partial f_i$ is nonempty-valued around $x_i$, then $\mathbf{x}$ is a lifted stationary point for (P). Conversely, if $\mathbf{x}$ is a lifted stationary point for (P) and, for each $i \in \{1, \ldots, m\}$, $f_i$ is strictly differentiable at $x_i$, then $(\mathbf{x}, \mathbf{y})$ is a stationary point for $(P_1)$.

Proof. (i): Since $F(\mathbf{x}) = h(\mathbf{x}) + \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)}$ with each $\frac{f_i}{g_i}$ Lipschitz continuous around $x_i$, it follows from the sum rule of the limiting subdifferential that

$$\partial_L (-F + \delta_S)(\mathbf{x}) \subseteq \partial_L (-h + \delta_S)(\mathbf{x}) + \partial_L \left( - \sum_{i=1}^m \frac{f_i}{g_i} \right)(\mathbf{x}). \tag{3.2}$$

Assume that $\mathbf{x}$ is a stationary point for (P) and, for each $i \in \{1, \ldots, m\}$, $\partial f_i$ is nonempty-valued around $x_i$. Then $0 \in \partial_L (-F + \delta_S)(\mathbf{x})$ and, by Lemma A.1(i)&(iv),

$$\partial_L \left( - \sum_{i=1}^m \frac{f_i}{g_i} \right)(\mathbf{x}) = \left( \begin{array}{c} \partial_L \left( - \frac{f_1}{g_1}(x_1) \right) \\ \vdots \\ \partial_L \left( - \frac{f_m}{g_m}(x_m) \right) \end{array} \right) \subseteq \left( \begin{array}{c} -g_1(x_1) \partial_L f_1(x_1) + f_1(x_1) \partial_L g_1(x_1) \\ \vdots \\ -g_m(x_m) \partial_L f_m(x_m) + f_m(x_m) \partial_L g_m(x_m) \end{array} \right), \tag{3.3}$$

which combined with (3.2) implies that $\mathbf{x}$ is a lifted stationary point for $(P_1)$.

Conversely, if $\mathbf{x}$ is a lifted stationary point and, for all $i \in \{1, \ldots, m\}$, $f_i, g_i$ are strictly differentiable at $x_i$, then the inclusions in (3.2) and (3.3) can be replaced by equalities, and hence $\mathbf{x}$ is a stationary point.

(ii): For each $i \in \{1, \ldots, m\}$, since $f_i(x_i) > 0$ and $\sqrt{g_i} \geq 0$, we have from Lemma A.1(ii), Lemma A.1(v), and then Lemma A.1(iii) that, if $\partial f_i$ is nonempty-valued around $x_i$, then

$$\partial^L_2(-H_i)(x_i, y_i) \subseteq \frac{y_i}{f_i(x_i)} \partial_L (-f_i)(x_i) + \frac{\sqrt{y_i}}{\sqrt{f_i(x_i)}} \partial_L g_i(x_i) \subseteq -\frac{\sqrt{y_i}}{\sqrt{f_i(x_i)}} \partial_L f_i(x_i) + \frac{y_i}{y_i^2} \partial_L g_i(x_i) = -g_i(x_i) \partial_L f_i(x_i) + f_i(x_i) \partial_L g_i(x_i). \tag{3.4}$$

Let $(\mathbf{x}, \mathbf{y})$ be a stationary point for $(P_1)$ and assume that, for each $i \in \{1, \ldots, m\}$, $\partial f_i$ is nonempty-valued around $x_i$. Then $0 \in \partial_L (-h + \delta_S)(\mathbf{x}) + \partial^L_2(-H)(\mathbf{x}, \mathbf{y})$ and (3.4) holds. Noting that $H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m H_i(x_i, y_i)$, one has $\partial^L_2(-H)(\mathbf{x}, \mathbf{y}) = \partial^L_2(-H_1)(x_1, y_1) \times \cdots \times \partial^L_2(-H_m)(x_m, y_m)$. Thus, $\mathbf{x}$ is a lifted stationary point for (P).

Conversely, let $\mathbf{x}$ be a lifted stationary point for (P) and assume that, for each $i \in \{1, \ldots, m\}$, $f_i$ is strictly differentiable at $x_i$. Then, the inclusions in (3.4) become equalities, and so $\mathbf{x}$ is a lifted stationary point for (P).
4. Inertial proximal block coordinate method

In this section, we propose an inertial proximal block coordinate method for solving the sum-of-ratios optimization problem \((P)\) and establish the convergence analysis for the proposed method. From now on, we will work under the following assumptions.

**Assumption A.** For problem \((P)\), \(S\) is a (not necessarily convex) closed set, \(-h\) is a proper lower semicontinuous function, and, for each \(i \in \{1, \ldots, m\}\), the functions \(f_i\) and \(g_i\) are locally Lipschitz functions on an open set containing \(S_i\). Moreover,

**A1.** For each \(i \in \{1, \ldots, m\}\), \(f_i\) is nonnegative on an open set containing \(S_i\) and there exists \(\alpha_i \geq 0\) such that, for all \(x_i, z_i \in S_i\) and all \(u \in \partial_L f_i(x_i)\),

\[
\frac{u}{2\sqrt{f_i(x_i)}} \leq \sqrt{f_i(z_i)} - \sqrt{f_i(x_i)} + \frac{\alpha_i}{2} \|z_i - x_i\|^2,
\]

whenever \(f_i(x_i) > 0\).

**A2.** For each \(i \in \{1, \ldots, m\}\), \(g_i\) is positive on \(S_i\) and there exists \(\beta_i \geq 0\) such that, for all \(x_i, z_i \in S_i\) and all \(v \in \partial_L g_i(x_i)\),

\[
\langle v, z_i - x_i \rangle \geq g_i(z_i) - g_i(x_i) - \frac{\beta_i}{2} \|z_i - x_i\|^2.
\]

**Remark 4.1 (Comments for the standing assumptions).** We note that standing Assumption A is quite general and, in particular, is satisfied for all the motivating examples.

(i) Assumption A1 is fulfilled if, for each \(i \in \{1, \ldots, m\}\), \(f_i\) takes nonnegative values on an open set \(O_i\) containing \(S_i\) and \(\sqrt{f_i}\) is weakly convex on \(O_i\) with modulus \(\alpha_i\). Clearly, this condition is true if \(f_i(x_i) = x_i^T A_i x_i\) for a positive semi-definite matrix \(A_i\) (as in the motivating Example 1.2 and Example 1.3) because \(\sqrt{f_i(x_i)} = \|A_i^{1/2} x_i\|\) which is convex, where \(A_i^{1/2}\) is a symmetric matrix such that \(A_i^{1/2} A_i^{1/2} = A_i\).

Assumption A1 also holds if, for each \(i \in \{1, \ldots, m\}\), \(S_i\) is compact, \(f_i\) takes positive values on an open set \(O_i\) containing \(S_i\) and \(f_i\) is a differentiable function whose gradient is Lipschitz continuous on \(O_i\) with modulus \(L_i\) for some \(L_i > 0\). Indeed, in this case, letting \(r_i := \min_{x_i \in S_i} f_i(x_i) > 0\), a direct verification shows that \(\sqrt{f_i}\) is weakly convex with modulus \(\alpha_i = \frac{r_i^2}{4} + \frac{1}{4} r_i^2 \max_{x_i \in S_i} \|\nabla f_i(x_i)\|^2\). This covers, in particular, the motivating Example 1.1, where \(f_i(x_i) = \log(1 + u^T x_i + r_i)\) for some \(u_i \in \mathbb{R}^d_+ \setminus \{0\}\) and \(r_i \geq 0\) and \(S_i = \{x_i \in \mathbb{R}^d_+ : x_i^{min} \leq x_i \leq x_i^{max}\}\) with \(0 < x_i^{min} \leq x_i^{max}\).

Similarly, Assumption A2 is satisfied if, for each \(i \in \{1, \ldots, m\}\), \(g_i\) is positive on \(S_i\) and it is a differentiable function whose gradient is Lipschitz continuous with modulus \(\beta_i\). Thus, combining these observations, we see that Assumption A is satisfied for all the motivating examples.

(ii) Finally, we also notice that the first condition in Assumption A1 ensures that, if \(x_i \in S_i\) and \(f_i(x_i) = 0\), then \(0 \in \partial_L f_i(x_i)\) for \(i \in \{1, \ldots, m\}\).

We now propose our inertial proximal block coordinate method for \((P)\). As we will see later on, this method can be seen as a proximal block coordinate method of Gauss-Seidel type applied to the equivalent non-convex problem \((P_1)\). It is worthwhile noting that, even when applied to the single-ratio case \((m = 1\) and \(h \equiv 0)\), our method here is totally different from the proximal type methods in [7, 8] which are based on Dinkelbach’s approach.

**Algorithm 1 (Inertial proximal block coordinate method).**

\begin{itemize}
  \item [\textbf{Step 1.}] Choose \(x_0 = x_0 = (x_{1,0}, \ldots, x_{m,0}) \in S\) and set \(n = 0\). Let \(\delta \in \mathbb{R}_+\) and \(\nu \in [0, \delta/2]\).
  \item [\textbf{Step 2.}] Set \(y_n = (y_{1,n}, \ldots, y_{m,n})\), where, for each \(i \in \{1, \ldots, m\}\),

\[
y_{i,n} = \frac{\sqrt{f_i(x_{i,n})}}{g_i(x_{i,n})}.
\]

Choose \(\tau_n \in \mathbb{R}\) such that \(\tau_n \geq \delta + \max_{i=1,\ldots,m} \left\{ \frac{1}{2} (2 y_{i,n} \alpha_i + y_{i,n}^2 \beta_i) \right\}\). Let \(\nu_n \in [0, \nu/\tau_n]\). For each \(i \in \{1, \ldots, m\}\),
\end{itemize}
set containing problem (P) is equivalent to (P can be interpreted as a proximal block coordinate method. To see this, we recall that, according to Lemma 3.1, So, the update for each

\[ w_{i,n} = \begin{cases} 
    y_{i,n} \frac{u_{i,n}}{\sqrt{f_i(x_{i,n})}} - y_i^2 v_{i,n} & \text{if } f_i(x_{i,n}) > 0, \\
    0 & \text{if } f_i(x_{i,n}) = 0.
\end{cases} \]

Denote \( h_{i,n+1}(x_i) := h(x_{i,n+1}, x_{i-1,n+1}, x_i, x_{i+1,n}, \ldots, x_{m,n}) \) and compute

\[ x_{i,n+1} = \arg\max_{x_i \in S_i} \left\{ h_{i,n+1}(x_i) - \tau_n \left\| x_i - z_{i,n} - \frac{1}{2\tau_n} w_{i,n} \right\|^2 \right\}. \]

Update \( x_{n+1} = (x_{1,n+1}, \ldots, x_{m,n+1}) \).

\textbf{Step 3.} If a termination criterion is not met, set \( n = n + 1 \) and go to Step 2.

Before establishing the convergence of Algorithm 1, we interpret it and comment on its computational costs.

**Remark 4.2 (Interpretation of Algorithm 1 as a proximal block coordinate methods).** Suppose that, for each \( i \in \{1, \ldots, m\} \), \( g_i \) is continuously differentiable on an open set containing \( S_i \). We will show that Algorithm 1 can be interpreted as a proximal block coordinate method. To see this, we recall that, according to Lemma 3.1, problem (P) is equivalent to (P1). As, for each \( i \in \{1, \ldots, m\} \), \( y_i \mapsto H_i(x_i, y_i) := 2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i) \) is a strongly concave one-variable quadratic function which admits a global maximizer at \( \frac{f_i(x_i)}{g_i(x_i)} \), one has

\[ y_{n+1} = \arg\max_{y \in \mathbb{R}^m} \left\{ h(x_{n+1}) + \sum_{i=1}^m H_i(x_{i,n+1}, y_i) \right\} = \arg\max_{y \in \mathbb{R}^m} \left\{ h(x_{n+1}) + H(x_{n+1}, y) \right\}. \]

Let \( i \in \{1, \ldots, m\} \). We see that, if \( f_i(x_{i,n}) > 0 \), then

\[ w_{i,n} = y_{i,n} \frac{u_{i,n}}{\sqrt{f_i(x_{i,n})}} - y_i^2 v_{i,n} \in y_{i,n}\mathcal{P}_L f_i(x_{i,n}) - y_i^2 \partial_L g_i(x_{i,n}) = \partial_L^2 H_i(x_{i,n}, y_{i,n}), \]

where last equality follows from the assumption that \( g_i \) is continuously differentiable on an open set containing \( S_i \). If \( f_i(x_{i,n}) = 0 \), then \( y_{i,n} = 0 \) and \( w_{i,n} = 0 \). Moreover, noting from Assumption A1 that \( \sqrt{f_i(x_i)} \geq 0 \) on an open set containing \( S_i \) and \( f_i(x_{i,n}) = 0 \) with \( x_{i,n} \in S_i \), one has \( 0 \in \partial_L (\sqrt{f_i}(x_{i,n})) \). This implies that, in the case when \( f_i(x_{i,n}) = 0 \), one also has

\[ w_{i,n} = 0 \in y_{i,n}\mathcal{P}_L (\sqrt{f_i}(x_{i,n})) - y_i^2 \partial_L g_i(x_{i,n}) = \partial_L^2 H_i(x_{i,n}, y_{i,n}). \]

So, the update for \( x_{n+1} = (x_{1,n+1}, \ldots, x_{m,n+1}) \) involves for \( i \in \{1, \ldots, m\} \)

\[ x_{i,n+1} = \arg\max_{x_i \in S_i} \left\{ h_{i,n+1}(x_i) - \tau_n \left\| x_i - \left( z_{i,n} + \frac{1}{2\tau_n} w_{i,n} \right) \right\|^2 \right\} \quad \text{with} \quad w_{i,n} \in \partial_L^2 H_i(x_{i,n}, y_{i,n}). \]

Combining the above observations, one sees that the proposed proximal algorithm can be regarded as a block coordinate proximal inertial subgradient algorithm applied to problem (P1): \( \max_{(x,y) \in S \times \mathbb{R}^m} \left\{ \varphi(x) + H(x, y) \right\} \), where proximal subgradient steps are applied cyclically to the \( x \)-variable, and a direct maximization step is applied to the \( y \)-variable.

**Remark 4.3 (Discussion on the computational costs).** The major computation cost lies in the update of \( x_{n+1} \) in Step 2. The update, for each \( i \in \{1, \ldots, m\} \),

\[ x_{i,n+1} = \arg\max_{x_i \in S_i} \left\{ h_{i,n+1}(x_i) - \tau_n \left\| x_i - \left( z_{i,n} + \frac{1}{2\tau_n} w_{i,n} \right) \right\|^2 \right\} \]

is equivalent to computing the proximal operator\(^2 \) of \( \frac{1}{\tau_n} (-h_{i,n+1} + \delta_{S_i}) \) at the point \( z_{i,n} + \frac{1}{\tau_n} w_{i,n} \). This can be done efficiently in many situations, for example, in the following cases:

\(^2\)The \textit{proximal operator} of a function \( \varphi \) at \( x \) is defined by \( \text{Prox}_\varphi(x) = \arg\min_y \left\{ \varphi(y) + \frac{1}{2\tau} \| y - x \|^2 \right\} \).
(i) if \( h \equiv 0 \), then this reduces to the projection onto the set \( S \), which, in many cases, has closed forms. This is the case when \( S \) is a box (as in the motivating Example 1.1), \( S \) is a sphere or a ball, \( S_i = \{ x : \|x\|_0 \leq r \} \) for \( r > 0 \), \( S_i = \{ x : \|x\| = 1 \} \) and \( \|x\|_0 \leq r \) (as in the motivation Example 1.3 of the sparse generalized eigenvalue problem with \( \phi(x) \) being the indicator function of the sparsity set) and \( S_i = \{ X \in \mathbb{R}^{p \times d} : X^T X = I_d \} \).

(ii) if \( m = 1 \), \( h(x) = -\lambda \|x\|_0 \) or \( h(x) = -\lambda \|x\|_1 \) with \( \lambda \geq 0 \), and \( S = \{ x : \|x\| = 1 \} \) (as in the motivating Example 1.3 of sparse generalized eigenvalue problem with \( \phi(x) \) being the cardinality regularization or \( l_1 \)-regularization), then the resulting proximal operator can be simplified to \( \text{argmin}_{x \in \mathbb{R}^d} \{ \|x + a\|^2 - \tau h(x) : \|x\| = 1 \} \) for some \( a \in \mathbb{R}^d \) and \( \tau \geq 0 \). This can be further rewritten as \( \text{argmin}_{x \in \mathbb{R}^d} \{ \langle 2a, x \rangle - \tau h(x) : \|x\| = 1 \} \), which has a closed form solution (see [23, Proposition 6] and [18]).

(iii) if \( h \) is a (possibly) nonconvex quadratic function and \( S_i = \{ x : \|x\| = 1 \} \) (as in the motivating Example 1.2), then the resulting problem is a nonconvex quadratic programming problem with norm constraint which is known as the trust region problem. In this case, this problem can be solved efficiently, for example, via semi-definite programming solvers.

(iv) if \( h \) can be expressed as finitely maximum of concave quadratic functions, that is, \( h(x) = \max_{1 \leq i \leq p} \{ \frac{1}{2} x^T A_i x + a_i^T x + \alpha_i \} \), where each \( A_i \) is negative semi-definite (and so, \( h_{i,n+1} \) can also be expressed in this form) and \( S_i \) is a polyhedral set, then this is equivalent to the solving of \( p \) many quadratic programming problems with linear inequality constraints, and so, it can be solved efficiently via quadratic programming solvers.

**Theorem 4.4 (Subsequential convergence).** Let \( (x_n)_{n \in \mathbb{N}} \) be the sequence generated by Algorithm 1. Suppose that Assumption A holds, that \( F \) is bounded from above on \( S \), and that the set \( \{ x \in S : F(x) \geq F(x_0) \} \) is bounded. Then the following statements hold:

(i) For all \( n \in \mathbb{N} \),

\[
F(x_n) - \nabla \|x_n - x_{n-1}\|^2 \leq F(x_{n+1}) - (\delta - \nabla)\|x_{n+1} - x_n\|^2.
\]

(ii) The sequence \( (x_n)_{n \in \mathbb{N}} \) is bounded and asymptotically regular. In particular,

\[
\sum_{n=0}^{+\infty} \|x_{n+1} - x_n\|^2 < +\infty.
\]

Moreover, the sequence \( (F(x_n))_{n \in \mathbb{N}} \) is convergent.

(iii) Suppose, in addition, that \( \limsup_{n \to +\infty} \tau_n = \tau < +\infty \) and that either \( m = 1 \) or \( h \) is continuous on \( S \cap \text{dom } h \). Then, for every cluster point \( x \) of \( (x_n)_{n \in \mathbb{N}} \), it holds that

\[
\lim_{n \to +\infty} F(x_n) = F(x)
\]

and \( x \) is a lifted stationary point for \( (P) \).

**Proof.** (i): From Step 2 of Algorithm 1, we have, for all \( n \geq 0 \) that \( x_n \in S \), \( y_{i,n} \geq 0 \), and \( x_{n+1} = (x_{1,n+1}, \ldots, x_{m,n+1}) \in S_1 \times \cdots \times S_m \), where, for each \( i \in \{1, \ldots, m\} \),

\[
x_{i,n+1} = \arg\max_{x_i \in S_i} \left\{ h_{i,n+1}(x_i) - \tau_n \|x_i - z_{i,n} - \frac{1}{2\tau_n} w_{i,n}\|^2 \right\}.
\]

Thus, for all \( i \in \{1, \ldots, m\} \) and all \( x_i \in S_i \),

\[
h_{i,n+1}(x_i) - \tau_n \|x_i - z_{i,n} - \frac{1}{2\tau_n} w_{i,n}\|^2 \leq h_{i,n+1}(x_{i,n+1}) - \tau_n \|x_{i,n+1} - z_{i,n} - \frac{1}{2\tau_n} w_{i,n}\|^2,
\]

and so

\[
h_{i,n+1}(x_{i,n}) - h_{i,n+1}(x_{i,n+1}) \leq -\tau_n \|x_{i,n+1} - z_{i,n}\|^2 + \tau_n \|x_i - z_{i,n}\|^2 + \langle w_{i,n}, x_{i,n+1} - x_i \rangle
\]

\[
= -\tau_n \|x_{i,n+1} - x_{i,n}\|^2 + \tau_n \|x_i - x_{i,n}\|^2 + \langle w_{i,n}, x_{i,n+1} - x_i \rangle
\]

\[
+ 2\tau_n \nu_n \langle x_{i,n+1} - x_i, x_{i,n} - x_{i,n-1} \rangle, \tag{4.1}
\]

9
where the last equality follows from the fact that \( z_{i,n} = x_{i,n} + \nu_n(x_{i,n} - x_{i,n-1}) \). Letting \( x_i = x_{i,n} \), we obtain

\[
h_{i,n+1}(x_{i,n}) - h_{i,n+1}(x_{i,n+1}) \leq -\tau_n\|x_{i,n+1} - x_{i,n}\|^2 + \langle w_{i,n}, x_{i,n+1} - x_{i,n} \rangle + 2\tau_n\nu_n\langle x_{i,n+1} - x_{i,n}, x_{i,n} - x_{i,n-1} \rangle.
\]

Since, for all \( i \in \{1, \ldots, m\} \),

\[
2\langle x_{i,n+1} - x_{i,n}, x_{i,n} - x_{i,n-1} \rangle \leq \|x_{i,n+1} - x_{i,n}\|^2 + \|x_{i,n} - x_{i,n-1}\|^2,
\]

it follows that

\[
h(x_n) - h(x_{n+1}) = \sum_{i=1}^{m} (h_{i,n+1}(x_{i,n}) - h_{i,n+1}(x_{i,n+1}))
\]

\[
\leq -\tau_n\|x_{n+1} - x_n\|^2 + \tau_n\nu_n\|x_n - x_{n-1}\|^2 + \sum_{i=1}^{m} \langle w_{i,n}, x_{i,n+1} - x_{i,n} \rangle,
\]

(4.2)

where the first equality follows from the definition of \( h_{i,n+1} \).

Next, we show that, for all \( i \in \{1, \ldots, m\} \) and all \( n \geq 0 \),

\[
\langle w_{i,n}, x_{i,n+1} - x_{i,n} \rangle \leq f_i(x_{i,n+1}) - \frac{f_i(x_{i,n})}{g_i(x_{i,n})} + (\tau_n - \delta)\|x_{i,n+1} - x_{i,n}\|^2.
\]

To see this, we consider first the case when \( f_i(x_{i,n}) > 0 \). Then

\[ w_{i,n} = 2y_{i,n} \frac{u_{i,n}}{\sqrt{f_i(x_{i,n})}} - y_{i,n}^2v_{i,n}. \]

Since \( u_{i,n} \in \partial_L f_i(x_{i,n}) \), the assumption on \( \sqrt{f_i} \) gives

\[ \left\langle \frac{u_{i,n}}{2\sqrt{f_i(x_{i,n})}}, x_{i,n+1} - x_{i,n} \right\rangle \leq \sqrt{f_i(x_{i,n+1})} - \sqrt{f_i(x_{i,n})} + \frac{\alpha_i}{2}\|x_{i,n+1} - x_{i,n}\|^2. \]

(4.3)

Since \( v_{i,n} \in \partial_L g_i(x_{i,n}) \), the assumption on \( g_i \) gives

\[ \langle v_{i,n}, x_{i,n+1} - x_{i,n} \rangle \geq g_i(x_{i,n+1}) - g_i(x_{i,n}) - \frac{\beta_i}{2}\|x_{i,n+1} - x_{i,n}\|^2. \]

(4.4)

Multiplying (4.3) by \( 2y_{i,n} \geq 0 \) and (4.4) by \( -y_{i,n}^2 \leq 0 \) and then adding them, and using \( H_i(x_i, y_i) := 2y_i\sqrt{f_i(x_i)} - y_i^2g_i(x_i) \), we obtain that

\[ \langle w_{i,n}, x_{i,n+1} - x_{i,n} \rangle \leq H_i(x_{i,n+1}, y_{i,n}) - H_i(x_{i,n}, y_{i,n}) + \frac{1}{2}(2y_{i,n}\alpha_i + g_{i,n}^2\beta_i)\|x_{i,n+1} - x_{i,n}\|^2. \]

(4.5)

On the other hand, if \( f_i(x_{i,n}) = 0 \), then \( w_{i,n} = 0 \), hence (4.5) still holds. In turn, from (4.5) and the fact that \( y_{i,n+1} \) is the maximizer of \( H_i(x_{i,n+1}, \cdot) \), we derive that, for all \( i \in \{1, \ldots, m\} \) and all \( n \geq 0 \),

\[ \langle w_{i,n}, x_{i,n+1} - x_{i,n} \rangle \leq \frac{f_i(x_{i,n+1})}{g_i(x_{i,n+1})} - \frac{f_i(x_{i,n})}{g_i(x_{i,n})} + (\tau_n - \delta)\|x_{i,n+1} - x_{i,n}\|^2, \]

where the last inequality follows by our choice of \( \tau_n \). Combining this with (4.2), we deduce that, for all \( n \geq 0 \),

\[ (\delta - \tau_n\nu_n)\|x_{n+1} - x_n\|^2 - \tau_n\nu_n\|x_n - x_{n-1}\|^2 \leq \left[ h(x_{n+1}) + \sum_{i=1}^{m} \frac{f_i(x_{i,n+1})}{g_i(x_{i,n+1})} \right] - \left[ h(x_n) + \sum_{i=1}^{m} \frac{f_i(x_{i,n})}{g_i(x_{i,n})} \right]. \]

Since \( \nu_n \leq \tau/\tau_n \), it holds that, for all \( n \geq 0 \),

\[ F(x_n) - \tau\|x_n - x_{n-1}\|^2 \leq F(x_{n+1}) - (\delta - \tau)\|x_{n+1} - x_n\|^2. \]
(ii): Let \( \theta_n := F(x_n) - \|x_n - x_{n-1}\|^2 \) for all \( n \geq 0 \). From the assumption that \( \tau \in [0, \delta/2) \), one has \( \delta - \tau > 0 \) and
\[
\theta_n \leq \theta_{n+1} - (\delta - 2\tau)\|x_{n+1} - x_n\|^2 \quad \forall n \geq 0.
\]  
(4.6)
This shows that \( (\theta_n)_{n \geq 0} \) is nondecreasing. As \( F \) is bounded from above on \( S \), there exists \( M > 0 \) such that \( \sup_{n \geq 0} F(x_n) \leq M. \) Then \( \sup_{n \geq 0} \theta_n \leq M, \) and so \( \theta_n \to \theta^* \) as \( n \to +\infty. \) Let \( k \geq 0. \) Summing (4.6) from \( n = 0 \) to \( k, \) we have
\[
(\delta - 2\tau) \sum_{n=0}^{k} \|x_{n+1} - x_n\|^2 \leq \theta_0 - \theta_k.
\]
Letting \( k \to +\infty, \) we see that \( \sum_{n=0}^{\infty} \|x_{n+1} - x_n\|^2 < +\infty. \) In particular, \( \|x_n\| \to 0 \) as \( n \to +\infty. \) Thus, \( F(x_n) \to \theta^* \) as \( n \to +\infty. \)

Next, to see the boundedness of \( (x_n)_{n \in \mathbb{N}}, \) we observe from the nondecreasing property of \( (\theta_n)_{n \geq 0} \) that
\[
F(x_n) \geq \theta_n \geq \theta_0 = F(x_0) - \|x_0 - x_{-1}\|^2 = F(x_0),
\]
where the last equality follows as \( x_{-1} = x_0. \) So, \( (x_n)_{n \in \mathbb{N}} \subseteq \{x \in S : F(x) \geq F(x_0)\}, \) and hence \( (x_n)_{n \in \mathbb{N}} \) is a bounded sequence by our assumption.

(iii): Let \( x \) be any cluster point of \( (x_n)_{n \in \mathbb{N}} \) and let \( (x_{k_n})_{n \in \mathbb{N}} \) be a subsequence of \( (x_n)_{n \in \mathbb{N}} \) such that \( x_{k_n} \to x = (\bar{x}_1, \ldots, \bar{x}_m) \) as \( n \to +\infty. \) Then \( x \in S \) and, by the asymptotic regularity, \( x_{k_n+1} \to x \) as \( n \to +\infty. \) This also shows that, for each \( i \in \{1, \ldots, m\}, \)
\[
\nu_{i,k_n} = \nu_{i,k_n} + \nu_{i,k_n}(x_{i,k_n} - x_{i,k_n+1}) \to \bar{x}_i \quad \text{as} \quad n \to +\infty.
\]
By the continuity of \( f_i \) and \( g_i, \) \( i \in \{1, \ldots, m\}, \) we have \( f_i(x_{k_n}) \to f_i(\bar{x}_i) \) and \( g_i(x_{k_n}) \to g_i(\bar{x}_i) > 0. \) Noting that \( f_i, g_i \) are locally Lipschitz and
\[
w_{i,k_n} = \begin{cases} \frac{g_i(x_{i,k_n})u_{i,k_n} - f_i(x_{i,k_n})v_{i,k_n}}{(g_i(x_{i,k_n}))^2} & \text{if } f_i(x_{i,k_n}) > 0, \\ 0 & \text{if } f_i(x_{i,k_n}) = 0, \end{cases}
\]
\[= \frac{g_i(x_{i,k_n})\partial_L f_i(x_{i,k_n}) - f_i(x_{i,k_n})\partial_L g_i(x_{i,k_n})}{(g_i(x_{i,k_n}))^2},\]
once sees that \( (w_{i,k_n})_{n \geq 0} \) is bounded, for each \( i \in \{1, \ldots, m\}. \) By passing to a subsequence if necessary, we can assume that, for each \( i \in \{1, \ldots, m\}, \)
\[
w_{i,k_n} \to w_i \in \frac{g_i(\bar{x}_i)\partial_L f_i(\bar{x}_i) - f_i(\bar{x}_i)\partial_L g_i(\bar{x}_i)}{g_i(\bar{x}_i)} \quad \text{as} \quad n \to +\infty.
\]
Set \( \bar{w} := (\bar{w}_1, \ldots, \bar{w}_m). \) We now split the proof in two following cases.

\textit{Case 1}: \( h \) is continuous on \( S \cap \text{dom } h. \) Then, \( h(x_{k_n}) \to h(x) \) as \( n \to +\infty. \) It follows that
\[
\lim_{n \to +\infty} F(x_n) = \lim_{n \to +\infty} F(x_{k_n}) = F(x)
\]
Replacing \( n \) by \( k_n \) in (4.1), we have, for all \( i \in \{1, \ldots, m\}, \) all \( x_i \in S_i \) and all \( n \geq 0, \)
\[
h_{i,k_n+1}(x_i) - h_{i,k_n+1}(x_{i,k_n}+1) \leq -\tau_k \|x_{i,k_n+1} - x_{i,k_n}\|^2 + \tau_{k_n} \|x_i - x_{i,k_n}\|^2 + (w_{i,n}, x_{i,k_n+1} - x_i)
\]  
\[+ 2\tau_{k_n} \nu_{i,k_n}(x_{i,k_n+1} - x_i, x_{i,k_n} - x_{i,k_n-1}). \]  
(4.7)
Letting \( n \to +\infty, \) one has
\[
h(\bar{x}_1, \ldots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \ldots, \bar{x}_m) - h(x) \leq \tau \|x_i - \bar{x}_i\|^2 + (\bar{w}_i, \bar{x}_i - x_i).
\]
This means that, for each \( i \in \{1, \ldots, m\}, \) the function
\[
\varphi_i(x_i) := -h(\bar{x}_1, \ldots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \ldots, \bar{x}_m) + \tau \|x_i - \bar{x}_i\|^2 - (\bar{w}_i, x_i) + \delta_S(\bar{x}_i)
\]
attains its minimum at \( \bar{x}_i, \) and so,
\[
0 \in \partial_L (-h + \delta_S)(\bar{x}_i) - \bar{w} \in \partial_L (-h + \delta_S)(\bar{x}_i) + \left( \begin{array}{c} -g_i(\bar{x}_i)\partial_L f_i(\bar{x}_i) + f_i(\bar{x}_i)\partial_L g_i(\bar{x}_i) \\ g_i(\bar{x}_i) \\ \vdots \\ -g_m(\bar{x}_m)\partial_L f_m(\bar{x}_m) + f_m(\bar{x}_m)\partial_L g_m(\bar{x}_m) \\ g_m(\bar{x}_m) \end{array} \right).
\]
Thus, $\mathbf{x}$ is a lifted stationary point for (P).

**Case 2:** $m = 1$. Then, (4.7) reduces to, for all $\mathbf{x} \in S$ and all $n \geq 0$,
\[
    h(\mathbf{x}) - h(\mathbf{x}_{k_n+1}) \leq -\tau_n \|\mathbf{x}_{k_n+1} - \mathbf{x}_{k_n}\|^2 + \tau_n \|\mathbf{x} - \mathbf{x}_{k_n}\|^2 + \langle \mathbf{w}_n, \mathbf{x}_{k_n+1} - \mathbf{x}\rangle + 2\tau_n \nu_{k_n}(\mathbf{x}_{k_n+1} - \mathbf{x}, \mathbf{x}_{k_n} - \mathbf{x}_{k_n-1}),
\]
where $\mathbf{w}_n \in \frac{g_1(\mathbf{x}_{k_n}) \partial f_1(\mathbf{x}_{k_n}) - f_1(\mathbf{x}_{k_n}) \partial g_1(\mathbf{x}_{k_n})}{g_1(\mathbf{x}_{k_n})^2}$ and $(\mathbf{w}_n)_{n \geq 0}$ is a bounded sequence. Letting $\mathbf{x} = \mathbf{x}$ and $n \to +\infty$, one has
\[
    \liminf_{n \to +\infty} h(\mathbf{x}_{k_n+1}) \geq h(\mathbf{x}).
\]
As $-h$ is lower semicontinuous, we have $\limsup_{n \to +\infty} h(\mathbf{x}_{k_n+1}) \leq h(\mathbf{x})$, and so, $h(\mathbf{x}_{k_n+1}) \to h(\mathbf{x})$ as $n \to +\infty$. It then follows that
\[
    \theta^* = \lim_{n \to +\infty} F(\mathbf{x}_n) = \lim_{n \to +\infty} F(\mathbf{x}_{k_n+1}) = F(\mathbf{x}).
\]
Then, by arguing as in **Case 1**, one sees that also in this case $\mathbf{x}$ is a lifted stationary point of (P).

**Remark 4.5 (Comments on the assumptions).** In addition to Assumption A, we also assume in Theorem 4.4 that the objective function $F$ is bounded from above on the feasible set $S$ and that \( \{ \mathbf{x} \in S : F(\mathbf{x}) \geq F(\mathbf{x}_0) \} \) is bounded. These assumptions are trivially satisfied in the case when $S$ is a compact set (as in our three motivating examples). More generally, they are also satisfied in the case when $-F$ is a coercive function on the set $S$ (noting that we are considering a maximization formulation), which is a standard assumption in the optimization literature.

Finally, in order to obtain that every cluster point is a lifted stationary point, we also assume that $\limsup_{n \to +\infty} \tau_n = \tau < +\infty$, and either $m = 1$ or $h$ is continuous on $S \cap \text{dom } h$. In the case $S$ is compact, the first assumption can be easily satisfied with $\tau_n = \delta + \max_{1 \leq i \leq m} \{ g_i(x_i, \alpha_i) + \frac{1}{2} \gamma_i^2 \}$ and $\tau = \delta + \max_{1 \leq i \leq m} \kappa_i$ with $\delta > 0$ and
\[
    \kappa_i = \frac{\sqrt{M_i}}{m_i} \alpha_i + \frac{M_i}{2m_i} \gamma_i,
\]
where $M_i := \max_{x_i \in S, f_i(x_i)}$ and $m_i := \min_{x_i \in S} g_i(x_i)$. For the second assumptions, the case $h$ is continuous on $S \cap \text{dom } h$ is quite general and it covers, in particular, the energy maximization problem (motivating Example 1.1) and the Rayleigh quotient optimization problem over unit sphere (motivating Example 1.2). Moreover, the case $m = 1$ allows us to cover cases such as the fractional quadratic optimization problems with sparse regularization (motivating Example 1.3).

**Remark 4.6 (Convergence to stronger stationary points).** A close inspection of the proof shows that one can obtain a stronger conclusion in Theorem 4.4(iii) for the cluster point $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_m)$. Indeed, the cluster point $\mathbf{x}$ satisfies the following stronger stationarity notion: for each $i \in \{1, \ldots, m\}$,
\[
    y_i \in \argmin_{x_i \in S_i} \{-h(\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, x_i, \mathbf{x}_{i+1}, \mathbf{x}_m) + \tau \|x_i - \mathbf{x}_i\|^2 - \langle \mathbf{w}_i, x_i \rangle\}
\]
for some $\mathbf{w}_i = \frac{g_i(\mathbf{x}) \partial f_i(\mathbf{x}) - f_i(\mathbf{x}) \partial g_i(\mathbf{x})}{g_i(\mathbf{x})^2}$. This relation implies that $\mathbf{x}$ is a lifted stationary point for (P). Moreover, in the case when $m = 1$, and $f_1, g_1$ are continuously differentiable, this relation reduces to
\[
    \mathbf{x} \in \text{Prox}_{\frac{1}{\tau}(-h + \delta_2)} \left( \mathbf{x} + \frac{1}{\tau} \nabla \left( \frac{f_1}{g_1} \right)(\mathbf{x}) \right),
\]
which corresponds to the notion of a $L$-stationary point with $L = 2\tau$ [4, Definition 4.8], a stronger notion than the usual one of a stationary point. This shows that Algorithm 1 can converge to a global solution for a nonsmooth and nonconvex fractional program with multiple stationary points, which will be illustrated in Example 6.1 later.

**Convergence of the full sequences**

Firstly, we fix some notation which will be used later on. Let $\mathcal{H}$ and $\mathcal{K}$ be two finite-dimensional real Hilbert spaces. Let $G : \mathcal{K} \to (-\infty, +\infty]$ be a proper lower semicontinuous function, $(\mathbf{x}_n)_{n \in \mathbb{N}}$ and $(\mathbf{z}_n)_{n \in \mathbb{N}}$ be sequences in $\mathcal{H}$ and $\mathcal{K}$, respectively, $(\alpha_n)_{n \in \mathbb{N}}$ and $(\beta_n)_{n \in \mathbb{N}}$ sequences in $\mathbb{R}_+$, $(\Delta_n)_{n \in \mathbb{N}}$ and $(\varepsilon_n)_{n \in \mathbb{N}}$ sequences in $\mathbb{R}_+$, and let $1 \leq \tau$ be two integers and $\lambda_i \in \mathbb{R}_+, i \in I := \{1, \frac{1}{2} + 1, \ldots, \tau\}$, with $\sum_{i \in I} \lambda_i = 1$. We set $\Delta_k = 0$ for $k \leq 0$ and consider the following conditions:
H1 (Sufficient decrease condition). For each $n \in \mathbb{N}$,
\[
G(z_{n+1}) + \alpha_n \Delta_n^2 \leq G(z_n);
\]

H2 (Relative error condition). For each $n \in \mathbb{N}$,
\[
\beta_n \text{dist}(0, \partial L G(z_n)) \leq \sum_{i \in I} \lambda_i \Delta_{n-i} + \varepsilon_n;
\]

H3 (Continuity condition). There exist a subsequence $(z_{k_n})_{n \in \mathbb{N}}$ and $\bar{z}$ such that
\[
z_{k_n} \to \bar{z} \quad \text{and} \quad G(z_{k_n}) \to G(\bar{z})\quad \text{as} \quad n \to +\infty;
\]

H4 (Parameter condition). It holds that
\[
\alpha := \inf_{n \in \mathbb{N}} \alpha_n > 0, \quad \gamma := \inf_{n \in \mathbb{N}} \alpha_n \beta_n > 0, \quad \text{and} \quad \sum_{n=1}^{+\infty} \varepsilon_n < +\infty;
\]

H5 (Distance condition). There exist $j \in \mathbb{Z}$ and $c \in \mathbb{R}$ such that, for all $n \in \mathbb{N}$,
\[
\|x_{n+1} - x_n\| \leq c \Delta_{n+j}.
\]

Theorem 4.7 (Abstract convergence [8]). Suppose that (H1), (H2), (H3), (H4) and (H5) hold and that the sequence $(z_n)_{n \in \mathbb{N}}$ is bounded. Let $\Omega$ be the set of cluster points of $(z_n)_{n \in \mathbb{N}}$ and suppose that $G$ is constant on $\Omega$ and satisfies the KL property at each point of $\Omega$. Then $\sum_{n=0}^{+\infty} \|x_{n+1} - x_n\| < +\infty$ and the sequence $(x_n)_{n \in \mathbb{N}}$ is convergent.

Theorem 4.8 (Global convergence). Let $(x_n)_{n \geq 0}$ be the sequence generated by Algorithm 1. Suppose that Assumption A holds, that, for each $i \in \{1, \ldots, m\}$, $f_i$ and $g_i$ are continuously differentiable on an open set containing $S_i$, that
\[
G(x, u) = -F(x) + \delta_S(x) + \nu \|x - u\|^2
\]
satisfies KL property at $(x, x)$ for all $x \in \text{dom} \partial L(-F + \delta_S)$, that $F$ is bounded from above on $S$, and that the set $\{x \in S : F(x) \geq F(x_0)\}$ is bounded. Suppose further that $\limsup_{n \to +\infty} \tau_n = \overline{\tau} < +\infty$, that either $m = 1$ or $h$ is a differentiable function on an open set containing $S$ whose gradient is Lipschitz continuous on $S$, and that there exist $\varepsilon, \ell \in \mathbb{R}_{++}$ satisfying
\[
\forall i \in \{1, \ldots, m\}, \forall x, x' \in S_i, \quad \|x - x'\| \leq \varepsilon \quad \implies \quad \|\nabla \left( \frac{f_i}{g_i} \right)(x) - \nabla \left( \frac{f_i}{g_i} \right)(x')\| \leq \ell \|x - x'\|.
\]

Then
\[
\sum_{n=0}^{+\infty} \|x_{n+1} - x_n\| < +\infty.
\]

In particular, the sequence $(x_n)_{n \geq 0}$ converges to a stationary point for $(P)$.

Proof. Let $z_n := (x_{n+1}, x_n)$ for all $n \geq 0$ and $\Omega$ be the set of cluster points of $(z_n)_{n \geq 0}$. We derive from Theorem 4.4 that the sequence $(z_n)_{n \geq 0} \subseteq S \times S$ is bounded and asymptotically regular, that, for all $n \geq 0$,
\[
G(z_{n+1}) + \alpha \|x_{n+2} - x_{n+1}\|^2 \leq G(z_n) \quad \text{with} \quad \alpha := \delta - 2\overline{\tau} > 0,
\]
and that, for every $z \in \Omega$, one has $\overline{z} = (\mathbf{x}, \mathbf{x})$ with $\mathbf{x} \in S$ such that $\mathbf{x}$ is a lifted stationary point and
\[
F(x_n) \to F(\mathbf{x}) \quad \text{as} \quad n \to +\infty.
\]

In particular, $\mathbf{x} \in \text{dom} \partial L(-F + \delta_S)$ and $G(z_n) = G(x_{n+1}, x_n) = -F(x_{n+1}) + \nu \|x_{n+1} - x_n\|^2 < -F(\mathbf{x})$ as $n \to +\infty$.

Now, as $\partial L G(z_n) = (\partial L(-F + \delta_S)(x_{n+1}) + 2\nu (x_{n+1} - x_n), 2\nu (x_n - x_{n+1}))^T$, it holds that
\[
\text{dist}(0, \partial L G(z_n)) = \sqrt{\text{dist}(0, \partial L(-F + \delta_S)(x_{n+1}) + 2\nu (x_{n+1} - x_n))^2 + (2\nu)^2 \|x_{n+1} - x_n\|^2}
\]

13
where the last inequality holds due to the assumption that $\partial L(-F + \delta_S)(x_{n+1}) = 0$. From Step 2 of Algorithm 1 and noting that $f_i, g_i$ are continuously differentiable on an open set that contains $S$, we have, for all $i \in \{1, \ldots, m\}$ and all $n \geq 0$, that

$$
0 \in \partial L(h_{i,n+1} + \delta_S)(x_{i,n+1}) + 2\tau_n(x_{i,n+1} - z_{i,n}) - w_{i,n} = \partial L \left( h_{i,n+1} - \frac{f_i}{g_i} + \delta_S \right) (x_{i,n+1}) + 2\tau_n(x_{i,n+1} - z_{i,n}) + (w_{i,n+1} - w_{i,n}),
$$

(4.11)

where

$$
w_{i,n} = \frac{g_i(x_{i,n})}{g_i(x_{i,n})}f_i(x_{i,n}) - f_i(x_{i,n})\frac{\nabla f_i(x_{i,n})}{g_i(x_{i,n})} = \nabla \left( \frac{f_i}{g_i} \right) (x_{i,n}).
$$

Since $\limsup_{n \to +\infty} \tau_n = \tau < +\infty$ and, for each $i \in \{1, \ldots, m\}$, $\lim_{n \to +\infty} \|x_{i,n+1} - x_{i,n}\| = 0$, there exists $n_0 \geq 0$ such that, for all $n \geq n_0$,

$$
\tau_n \leq 2\tau \quad \text{and} \quad \|x_{i,n+1} - x_{i,n}\| \leq \varepsilon. \quad \text{(4.12)}
$$

Then, for all $i \in \{1, \ldots, m\}$ and all $n \geq n_0$, we derive from $\nu_n \leq \tau / \tau_n$ that

$$
\|2\tau_n(x_{i,n+1} - z_{i,n})\| = \|2\tau_n(x_{i,n+1} - x_{i,n}) - 2\tau_n\nu_n(x_{i,n} - x_{i,n-1})\|
\leq 4\tau\|x_{i,n+1} - x_{i,n}\| + 2\tau\|x_{i,n} - x_{i,n-1}\|
$$

(4.12)

and from the assumption on $\nabla \left( \frac{f_i}{g_i} \right)$ that

$$
\|w_{i,n+1} - w_{i,n}\| = \left\| \nabla \left( \frac{f_i}{g_i} \right) (x_{i,n+1}) - \frac{\nabla (f_i)}{g_i} (x_{i,n}) \right\| \leq \ell\|x_{i,n+1} - x_{i,n}\|. \quad \text{(4.13)}
$$

We split the discussion into the following cases.

Case 1: $h$ is a differentiable function on an open set containing $S$ whose gradient is Lipschitz continuous on $S$ with modulus $\ell_h$. Then, it follows from (4.11) that, for all $i \in \{1, \ldots, m\}$ and all $n \geq 0$,

$$
0 \in -\nabla h_{i,n+1}(x_{i,n+1}) + \partial L \left( \frac{-f_i}{g_i} + \delta_S \right) (x_{i,n+1}) + 2\tau_n(x_{i,n+1} - z_{i,n}) + (w_{i,n+1} - w_{i,n}),
$$

which yields

$$
-\|\nabla h(x_{n+1}) - \nabla h_{i,n+1}(x_{i,n+1})\| - 2\tau_n(x_{i,n+1} - z_{i,n}) - (w_{i,n+1} - w_{i,n}) \in \partial L(-F + \delta_S)(x_{n+1}).
$$

Combining with (4.12) and (4.13), we deduce that for all $n \geq n_0$,

$$
\text{dist}(0, \partial L(-F + \delta_S)(x_{n+1})) \leq \sum_{i=1}^m \text{dist}(0, \partial L^i(-F + \delta_S)(x_{n+1}))
\leq \sum_{i=1}^m \|\nabla h(x_{n+1}) - \nabla h_{i,n+1}(x_{i,n+1})\| + (4\tau + \ell) \sum_{i=1}^m \|x_{i,n+1} - x_{i,n}\|
\quad + 2\tau \sum_{i=1}^m \|x_{i,n} - x_{i,n-1}\|
\leq m\ell_h \|x_{n+1} - x_n\| + (4\tau + \ell)\sqrt{m}\|x_{n+1} - x_n\| + 2\sqrt{m}\|x_n - x_{n-1}\|,
$$

where the last inequality holds due to the assumption that $\nabla h$ is Lipschitz continuous with modulus $\ell_h$ on $S$. By using (4.10), there exists $K \in \mathbb{R}_{++}$ such that, for all $n \geq n_0$,

$$
\text{dist}(0, \partial L G(z_n)) \leq K \left( \|x_{n+1} - x_n\| + \|x_n - x_{n-1}\| \right).
$$

The conclusion then follows by applying Theorem 4.7 with $I = \{1, 2\}$, $\lambda_1 = \lambda_2 = 1/2$, $\Delta_n = 2 K \|x_{n+2} - x_{n+1}\|$, $\alpha_n \equiv \frac{\sqrt{2}}{4\tau} > 0$, $\beta_n \equiv 1$, and $\varepsilon_n \equiv 0$. 

14
Case 2: $m = 1$. In this case, we derive from (4.11) that, for all $n \geq 0$,

$$-2r_n(x_{n+1} - z_n) - (w_{n+1} - w_n) \in \partial_L(-F + \delta_S)(x_{n+1}).$$

Thus, (4.12) and (4.13) imply that, for all $n \geq n_0$,

$$\text{dist}(0, \partial_L(-F + \delta_S)(x_{n+1})) \leq \|2r_n(x_{n+1} - z_n)\| + \|w_{n+1} - w_n\|$$

$$\leq (4\tau + \ell)\|x_{n+1} - x_n\| + 2\tau\|x_n - x_{n-1}\|.$$

Proceeding as in Case 1, we obtain the desired conclusion. ■

Finally, we remark that, if the KL exponent $\alpha$ of the merit function $G$ is available, then the following proposition on the convergence rate of the algorithm holds. As the techniques are rather standard, we omit the proof here and refer the readers to [1, 16].

Proposition 4.9. Suppose that the assumptions of Theorem 4.8 hold and that the KL exponent of $G$ at $(\bar{x}, \tilde{x})$ is $\alpha \in [0,1)$ for all $\bar{x} \in \text{dom} \partial_L(-F + \delta_S)$. Let $(x_n)_{n \geq 0}$ be the sequence generated by Algorithm 1. Then the following statements hold:

(i) if $\alpha = 0$, then $(x_n)_{n \geq 0}$ converges after finitely many iterations;

(ii) if $\alpha \in (0, \frac{1}{2}]$, then $(x_n)_{n \geq 0}$ converges to $\bar{x}$ linearly for some stationary point $\bar{x}$ for $(P)$, that is, there exist $r \in (0, 1)$ and $M > 0$ such that, for all $n \geq 0$, $\|x_n - \bar{x}\| \leq Mr^n$;

(iii) if $\alpha \in (\frac{1}{2}, 1)$, then $(x_n)$ converges to $\bar{x}$ sublinearly with rate $O(n^{-\frac{1-(n-\alpha)}{2n-1}})$ for some stationary point $\bar{x}$ for $(P)$,

that is, there exists $M > 0$ such that, for all $n \geq 0$, $\|x_n - \bar{x}\| \leq Mn^{-\frac{1-(n-\alpha)}{2n-1}}$.

As stated in the preceding corollary, the KL exponent of the merit function for the model problem completely determines the convergence rate of the proposed algorithm. On the other hand, finding or estimating the KL exponent of a nonsmooth and nonconvex function is, in general, highly challenging. In the next section, we will derive KL exponents of the corresponding merit functions for various classes of structured fractional programming problems.

5. KL exponents for structured fractional programs

In this section, we derive the KL exponent of the associated merit functions of three classes of structured fractional programs: sum-of-ratios fractional quadratic programs with spherical constraints, generalized eigenvalue problems with cardinality regularization and generalized eigenvalue problems with sparsity constraints. In particular, we establish that, for the last two classes of fractional programs, the KL exponent is $1/2$. As a consequence, the proposed Algorithm 1 exhibits linear convergence for these two classes of fractional programs.

We first see that the KL exponent for the merit function associated with $(P)$ can be computed by a merit function associated with the equivalent problem $(P_1)$. To do this, we need the following result from [16].

Lemma 5.1 (cf. [16, Theorem 3.6]). Let $f$ be a proper lower semicontinuous function. Suppose that $f$ satisfies the KL property at $x \in \text{dom} \partial_L f$ with KL exponent $\alpha \in [\frac{1}{2}, 1)$. Then, for all $\rho \geq 0$, $f(x, u) = f(x) + \rho\|x - u\|^2$ satisfies the KL property with KL exponent $\alpha$ at $(x, x)$.

Proposition 5.2. Suppose that Assumption A holds and that, for each $i \in \{1, \ldots, m\}$, $f_i$ and $g_i$ are continuously differentiable on $S_i$. Let $P(x, y) = -h(x) - H(x, y) + \delta_S(x)$, where $H(x, y) = \sum_{i=1}^m \left[2y_i \sqrt{f_i(x_i)} - y_i^2 g_i(x_i)\right]$. Let $x = (x_1, \ldots, x_m) \in S$ and $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$ with $y_i = \frac{\sqrt{f_i(x_i)}}{g_i(x_i)}$. Suppose further that $h$ is continuous around $x$ and that $P$ satisfies the KL property with KL exponent $\alpha \in [0, 1)$ at $(x, y) \in S \times \mathbb{R}^m$. Then

$$\Phi(x) := -h(x) - \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)} + \delta_S(x)$$

satisfies the KL property with KL exponent $\alpha$ at $x$. In particular, for all $\rho \geq 0$,

$$G(x, u) := -h(x) - \sum_{i=1}^m \frac{f_i(x_i)}{g_i(x_i)} + \rho\|x - u\|^2 + \delta_S(x)$$

15
satisfies the KL property with KL exponent $\alpha' = \max\{\alpha, \frac{1}{2}\}$ at $(\mathbf{x}, \mathbf{x})$.

Proof. As $P$ satisfies the KL property with KL exponent $\alpha \in [0, 1)$ at $(\mathbf{x}, \mathbf{y}) \in \mathcal{H} \times \mathbb{R}^m$, there exist $\delta, \eta, c > 0$ such that, for all $(\mathbf{x}, \mathbf{y})$ with $\|\mathbf{x} - \mathbf{y}\| \leq \delta$ and $P(\mathbf{x}, \mathbf{y}) < P(\mathbf{x}, \mathbf{y}) < P(\mathbf{x}, \mathbf{y}) + \eta$,

$$\text{dist}(0, \partial_L P(\mathbf{x}, \mathbf{y})) \geq c [P(\mathbf{x}, \mathbf{y}) - P(\mathbf{x}, \mathbf{y})] \alpha. \quad (5.1)$$

It follows that, for all $(\mathbf{x}, \mathbf{y})$ with $\|\mathbf{x} - \mathbf{y}\| \leq \delta$ and $P(\mathbf{x}, \mathbf{y}) < P(\mathbf{x}, \mathbf{y}) + \eta$,

$$[\text{dist}(0, \partial_L P(\mathbf{x}, \mathbf{y}))]^{\frac{1}{\alpha}} \geq c^{\frac{1}{\alpha}} [P(\mathbf{x}, \mathbf{y}) - P(\mathbf{x}, \mathbf{y})]. \quad (5.2)$$

Here, we drop the condition $P(\mathbf{x}, \mathbf{y}) < P(\mathbf{x}, \mathbf{y})$ because (5.2) trivially holds otherwise. For each $\mathbf{x}$, let $\mathbf{y}_\mathbf{x} = (y_{1, \mathbf{x}}, \ldots, y_{m, \mathbf{x}})$ with $y_{i, \mathbf{x}} = \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})}$ for $i \in \{1, \ldots, m\}$.

Therefore, from (5.2) we derive that, for all $\mathbf{x} \in S$ with $\|\mathbf{x} - \mathbf{x}\| \leq \delta$,

$$[\text{dist}(0, \partial_L P(\mathbf{x}, \mathbf{y}_\mathbf{x}))]^{\frac{1}{\alpha}} \geq c^{\frac{1}{\alpha}} [h(\mathbf{x}) - H(\mathbf{x}, \mathbf{y}_\mathbf{x}) + h(\mathbf{x}) + H(\mathbf{x}, \mathbf{y})]$$

$$= c^{-\frac{1}{\alpha}} \left[-h(\mathbf{x}) - \sum_{i=1}^{m} \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} + \sum_{i=1}^{m} \frac{f_i(\mathbf{x})}{g_i(\mathbf{x})} \right].$$

Now, we notice that $\partial_L P(\mathbf{x}, \mathbf{y}_\mathbf{x}) = (\partial_L (-h + \delta \mathbf{S}))(\mathbf{x}) + \partial_L^2 (-H)(\mathbf{x}, \mathbf{y}_\mathbf{x}), \partial_L^2 (-H)(\mathbf{x}, \mathbf{y}_\mathbf{x})$ and that, for each $i \in \{1, \ldots, m\}$,

$$\partial_L^2 (-H_i)(x_i, y_{i, \mathbf{x}}) = \frac{-g_i(x_i) \nabla f_i(x_i) + f_i(x_i) \nabla g_i(x_i)}{[g_i(x_i)]^2} = \nabla \left( \frac{f_i}{g_i} \right)(x_i), \text{ and } \partial_L^2 (-H_i)(x_i, y_{i, \mathbf{x}}) = 0.$$  

Therefore, $\partial_L P(\mathbf{x}, \mathbf{y}_\mathbf{x}) = (\partial_L \Phi(\mathbf{x}), 0)$, and from here we deduce that, for all $\mathbf{x} \in S$ with $\|\mathbf{x} - \mathbf{x}\| \leq \delta$,

$$[\text{dist}(0, \partial_L \Phi(\mathbf{x}))]^{\frac{1}{\alpha}} \geq c^{\frac{1}{\alpha}} [\Phi(\mathbf{x}) - \Phi(\mathbf{x})].$$

So, $\Phi$ satisfies the KL property with exponent $\alpha$ at $\mathbf{x}$, and hence, also with KL exponent $\max\{\alpha, \frac{1}{2}\}$. By using Lemma 5.1, $G$ satisfies the KL property with exponent $\max\{\alpha, \frac{1}{2}\}$ at $(\mathbf{x}, \mathbf{x})$. \hfill \Box

### 5.1. Sum-of-ratios fractional quadratic programs with spherical constraint

We now consider the following sum-of-ratios fractional quadratic program

$$\max_{\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}^d} \mathbf{x}^\top A_0 \mathbf{x} + a_0^\top \mathbf{x} + \sum_{i=1}^{m} \frac{x_i^\top A_i x_i}{x_i^\top B_i x_i} \text{ s.t. } \|x_i\| = 1, \quad i \in \{1, \ldots, m\}, \quad (\text{FQP})$$

where $A_i$ and $B_i$ are positive definite matrices. In the special cases of $m = 1$ and $a_0 = 0$, this reduces to the problem of maximizing the sum of a quadratic function and the Rayleigh quotient over the unit sphere (motivating Example 1.2). For this sum-of-ratios fractional quadratic program, the corresponding merit function for the proposed Algorithm 1 takes the form

$$\hat{\Phi}_{\text{FQP}}(\mathbf{x}, \mathbf{u}) = - [\mathbf{x}^\top A_0 \mathbf{x} + a_0^\top \mathbf{x}] - \sum_{i=1}^{m} \frac{x_i^\top A_i x_i}{x_i^\top B_i x_i} + \delta \Lambda_1 \times \cdots \times \Lambda_m (\mathbf{x}) + \rho \|\mathbf{x} - \mathbf{u}\|^2,$$

where $\Lambda_i = \{x_i \in \mathbb{R}^{d_i} : \|x_i\| = 1\}, \quad i \in \{1, \ldots, m\}$, and $\rho \geq 0$. Below, we investigate the KL exponent of this merit function.

To this end we will use a fundamental result which provides an exponent estimate in the classical Łojasiewicz gradient inequality for polynomials.

**Lemma 5.3 (Łojasiewicz gradient inequality [10, Theorem 4.2]).** Let $f$ be a polynomial on $\mathbb{R}^d$ with degree $p \in \mathbb{N}$. Suppose that $f(\mathbf{x}) = 0$. Then there exist constants $\varepsilon, c > 0$ such that, for all $\mathbf{x} \in \mathbb{R}^d$ with $\|\mathbf{x} - \mathbf{x}\| \leq \varepsilon$, we have

$$\|\nabla f(\mathbf{x})\| \geq c |f(\mathbf{x})|^{1-\tau}, \quad \text{where } \tau = \mathcal{R}(d, p)^{-1} \text{ and } \mathcal{R}(d, p) := \begin{cases} 1 & \text{if } p = 1, \\ p(3p - 3)^{d-1} & \text{if } p \geq 2. \end{cases} \quad (5.3)$$
Theorem 5.4. Let $\Lambda = \Lambda_1 \times \cdots \times \Lambda_m$, where $\Lambda_i = \{x_i \in \mathbb{R}^{d_i} : \|x_i\| = 1\}, i \in \{1, \ldots, m\}$, with $\sum_{i=1}^{m} d_i = d$. Consider

$$\Phi(x) = -[x^T A_0 x + a_0^T x] - \sum_{i=1}^{m} x_i^T A_i x_i + \delta(x),$$

where $A_i$ and $B_i$ are positive definite matrices. Then $\Phi$ satisfies the KL property with KL exponent $1 - \tau$ for all $\rho \geq 0$,

$$\Phi_{FQ}(x, u) = -[x^T A_0 x + a_0^T x] - \sum_{i=1}^{m} x_i^T A_i x_i + \delta(x) + \rho\|x - u\|^2$$

satisfies the KL property with KL exponent $1 - \tau$ at $(x, x)$ for all $x \in \text{dom} \partial \Phi$.

Proof. From Proposition 5.2 with $S = \Lambda$, $h(x) = x^T A_0 x + a_0^T x$, $f_i(x_i) = x_i^T A_i x_i$, and $g_i(x_i) = x_i^T B_i x_i, i \in \{1, \ldots, m\}$, it suffices to show that

$$P(x, y) = -[x^T A_0 x + a_0^T x] - \sum_{i=1}^{m} \left[ y_i \|L_i x_i\| - y_i^2 x_i^T B_i x_i \right] + \sum_{i=1}^{m} \delta_i(x_i).$$

As $A_i$ is positive definite, we have $L_i x_i \neq 0$ for all $x_i \in \Lambda_i$ and all $i \in \{1, \ldots, m\}$. Let $f_0(x) = x^T A_0 x + a_0^T x$. Then, for all $i \in \{1, \ldots, m\}$,

$$\partial_{x_i}^2 P(x, y) = \left\{ -\nabla_{x_i} f_0(x) - y_i \frac{A_i x_i}{\sqrt{x_i} A_i x_i} + 2y_i^2 B_i x_i + t_i x_i : t_i \in \mathbb{R} \right\}$$

and

$$\partial_{y_i}^2 P(x, y) = -\|L_i x_i\| + 2y_i x_i^T B_i x_i,$$

which imply that

$$\text{dist}(0, \partial_L P(x, y))^2 = \sum_{i=1}^{m} \inf_{t_i \in \mathbb{R}} \left\{ \| -\nabla_{x_i} f_0(x) - y_i \frac{A_i x_i}{\sqrt{x_i} A_i x_i} + 2y_i^2 B_i x_i + t_i x_i \|^2 \right\} + \sum_{i=1}^{m} (-\|L_i x_i\| + 2y_i x_i^T B_i x_i)^2.$$
and let $\hat{f} = f - r$ where $r = f(x, y, u, \bar{x}, \mu)$ with $u_i = \frac{L_i x_i}{\|L_i x_i\|}$, $\bar{x}_i = \frac{y_i}{\|L_i x_i\|}$ and $\mu_i = \frac{t_i x_i}{y_i}$, for all $i \in \{1, \ldots, m\}$.

Clearly, $\hat{f}$ is a polynomial on $\mathbb{R}^{d+3m+m}$ of degree 4. So, Lemma 5.3 implies that there exist $\delta_0 > 0$ and $c > 0$ such that, for all $(x, y, u, \lambda, \mu)$ with $\|(x, y, u, \lambda, \mu) - (x, y, u, \bar{x}, \mu)\| \leq \delta_0$,

$$\|\nabla f(x, y, u, \lambda, \mu)\| = \|\nabla \hat{f}(x, y, u, \lambda, \mu)\| \geq c \|\nabla \hat{f}(x, y, u, \lambda, \mu)\|^{1-\tau} = c |f(x, y, u, \lambda, \mu) - f(x, y, u, \bar{x}, \mu)|^{1-\tau},$$

where $\tau \geq (\mathcal{R}(d + 3m + md, 4))^{-1}$. Let $u_x = (u_{1,x}, \ldots, u_{m,x})$, $\lambda_{x,y} = (\lambda_{1,x,y}, \ldots, \lambda_{m,x,y})$, and $\mu_{x,y} = (\mu_{1,x,y}, \ldots, \mu_{m,x,y})$ with

$$u_{i,x} = \frac{L_i x_i}{\|L_i x_i\|}, \quad \lambda_{i,x,y} = \frac{y_i \|L_i x_i\|}{2}, \quad \text{and} \quad \mu_{i,x,y} = \frac{t_i x_i}{y_i}$$

for all $i \in \{1, \ldots, m\}$.

By shrinking $\delta > 0$ if necessary, we can assume that, for all $(x, y) \in \mathcal{A} \times \mathbb{R}^m$ with $\|(x, y) - (x, y)\| \leq \delta$,

$$\|(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) - (x, y, u_x, \bar{x}, \mu)\| \leq \delta_0,$$

and so,

$$\|\nabla f(x, y, u_x, \lambda_{x,y}, \mu_{x,y})\| \geq \|f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) - f(x, y, u_x, \bar{x}, \mu)\|^{1-\tau}.$$ 

Note that, for all $i \in \{1, \ldots, m\}$,

$$\left\{ \begin{array}{l}
\nabla_{x,i} f(x, y, u, \lambda, \mu) = -\nabla_{x,i} f_0(x) - (y_i L_i^\top x_i - 2y_i^2 B_i x_i) + 2\mu_i x_i, \\
\nabla_{y,i} f(x, y, u, \lambda, \mu) = -(L_i x_i)^\top u_i + 2y_i x_i^\top B_i x_i, \\
\nabla_{u,i} f(x, y, u, \lambda, \mu) = -y_i L_i x_i + 2\lambda_i u_i, \\
\nabla_{\lambda,i} f(x, y, u, \lambda, \mu) = \|u_i\|^2 - 1, \\
\nabla_{\mu,i} f(x, y, u, \lambda, \mu) = \|x_i\|^2 - 1.
\end{array} \right.$$ 

Direct verification shows that, for all $(x, y) \in \mathcal{A} \times \mathbb{R}^m$ with $\|(x, y) - (x, y)\| \leq \delta$ and all $i \in \{1, \ldots, m\}$, one has

$$\left\{ \begin{array}{l}
\nabla_{x,i} f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) = -\nabla_{x,i} f(x) - \left(\frac{y_i \lambda_{x,y}}{\sqrt{x_i^\top A_{x,y}}x_i} - 2y_i^2 B_i x_i\right) + t_i x_i, \\
\nabla_{y,i} f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) = -(L_i x_i)^\top u_i + 2y_i x_i^\top B_i x_i, \\
\nabla_{u,i} f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) = 0, \\
\nabla_{\lambda,i} f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) = 0, \\
\nabla_{\mu,i} f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) = 0
\end{array} \right.$$ 

and also

$$f(x, y, u_x, \lambda_{x,y}, \mu_{x,y}) = P(x, y) \quad \text{and} \quad f(\bar{x}, y, u_x, \bar{x}, \mu) = P(\bar{x}, y).$$

These together with (5.4) implies that, for all $(x, y) \in \mathcal{A} \times \mathbb{R}^m$ with $\|(x, y) - (x, y)\| \leq \delta$ and $P(x, y) < P(x, y) < P(\bar{x}, y) + \eta$,

$$\text{dist}(0, \partial L P(x, y)) \geq |P(x, y) - P(\bar{x}, y)|^{1-\tau} = [P(x, y) - P(\bar{x}, y)]^{1-\tau}.$$ 

Thus, $P$ satisfies the KL property with exponent $1 - \tau$, and the conclusion follows. \hfill \blacksquare

### 5.2. Generalized eigenvalue problem with cardinality regularization

Consider the generalized eigenvalue problem with cardinality regularization

$$\max_{x \in \mathbb{R}^d} \frac{x^\top A_1 x - \lambda \|x\|_0}{x^\top B_1 x} \quad \text{s.t.} \quad \|x\| = 1,$$

where $A_1, B_1$ are symmetric matrices such that $A_1$ is positive semidefinite and $B_1$ is positive definite, and $\lambda > 0$.

For this generalized eigenvalue problem with cardinality regularization the corresponding merit function for the proposed Algorithm 1 takes the form

$$\hat{\Phi}_{GEP}(x, u) = \frac{x^\top A x}{x^\top B x} + \lambda \|x\|_0 + \delta_1(x) + \rho \|x - u\|^2,$$

with $A = -A_1$ a symmetric matrix, $B = B_1$ is a positive definite matrix, $\Lambda = \{x \in \mathbb{R}^d : \|x\| = 1\}$, and $\rho \geq 0$.

Below, we derive the KL exponent of the merit function $\hat{\Phi}_{GEP}$.

To this end, we will use the following lemma from [16]. Here we provide an alternative short proof for it.
Lemma 5.5. Let $Q$ be a symmetric $d \times d$ matrix. Then, there exists $c > 0$ such that, for all $x \in \mathbb{R}^d$,
$$
\|Qx\|^2 \geq c (x^T Q x).
$$

Proof. Let $Q = U^T \Sigma U$ where $U$ is an orthonormal matrix and $\Sigma = \text{diag}(\lambda_1, \ldots, \lambda_n)$ is a diagonal matrix whose diagonal elements are the eigenvalues of $Q$ with $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Let $x \in \mathbb{R}^d$, $y := Ux$, and $I_0 = \{j : \lambda_j \neq 0\}$. Then $x^T Q x = \sum_{j=1}^N \lambda_j y_j^2 = \sum_{j \in I_0} \lambda_j y_j^2$ and
$$
\|Qx\|^2 = x^T (Q^T Q) x = (Ux)^T \Sigma^2 (Ux) = \sum_{j=1}^N \lambda_j^2 y_j^2 = \sum_{j \in I_0} \lambda_j^2 y_j^2.
$$

Let $c = \min\{|\lambda_j| : j \in I_0\}$. A direct verification shows that
$$
c (x^T Q x) = \min\{|\lambda_j| : j \in I_0\} \left( \sum_{j \in I_0} \lambda_j y_j^2 \right) \leq \min\{|\lambda_j| : j \in I_0\} \left( \sum_{j \in I_0} |\lambda_j| y_j^2 \right) \leq \sum_{j \in I_0} \lambda_j^2 y_j^2 = \|Qx\|^2.
$$

Thus, the conclusion follows.

Next we prove that the KL exponent of the merit function $\hat{\Phi}_{GEP}$ is $\frac{1}{2}$. To do this, for an index set $J = \{j_1, \ldots, j_k\} \subseteq \{1, \ldots, d\}$ with $k \leq d$, we denote $x_J := (x_{j_1}, \ldots, x_{j_k})$. Moreover, for two index sets $I, J$, we denote $A_{I,J} = (A_{ij})_{i \in I, j \in J}$.

Theorem 5.6. Consider the function
$$
\Phi(x) = \frac{x^T A x}{x^T B x} + \lambda \|x\|_0 + \delta_\Lambda(x),
$$
where $\Lambda = \{x : \|x\|_0 = 1\}$, $A, B$ are symmetric matrices with $B$ positive definite, and $\lambda > 0$. Then $\Phi$ is a KL function with exponent $\frac{1}{2}$. In particular, for all $\rho \geq 0$,
$$
\hat{\Phi}_{GEP}(x, u) = \frac{x^T A x}{x^T B x} + \lambda \|x\|_0 + \delta_\Lambda(x) + \rho \|x - u\|^2
$$
satisfies the KL property with KL exponent $\frac{1}{2}$ at $(\bar{x}, \bar{x})$ for all $\bar{x} \in \text{dom} \partial L \Phi$.

Proof. Take any $\bar{x} \in \text{dom} \partial L \Phi$. Then $\bar{x} \in \Lambda$. Let $J = \text{supp}(\bar{x})$ and use $|J|$ to denote the cardinality of $J$. Choose $\eta \in (0, 1)$ such that, for all $\|x - \bar{x}\| < \eta$,
$$
\left| \frac{x^T A x}{x^T B x} - \frac{\bar{x}^T A \bar{x}}{\bar{x}^T B \bar{x}} \right| \leq \frac{\lambda}{4} \text{ and } \eta < \frac{\lambda}{4}.
$$

Let $x$ with $\|x - \bar{x}\| < \eta$ and $\Phi(x) < \Phi(\bar{x}) < \Phi(\bar{x}) + \eta$. We first see that, by shrinking $\eta$ if necessary, one can assume that
$$
J = \text{supp}(x) = \text{supp}(x).
$$
Indeed, by continuity and by shrinking $\eta$ if necessary, one has $\text{supp}(x) \subseteq \text{supp}(\bar{x})$. Suppose that $\text{supp}(x) \subset \text{supp}(\bar{x})$. Then $\|x\|_0 > \|\bar{x}\|_0$, and so, $\|x\|_0 \geq \|\bar{x}\|_0 + 1$. From our choice of $x$, one has $x \in \Lambda$ and
$$
\frac{x^T A x}{x^T B x} + \lambda \|x\|_0 < \frac{\bar{x}^T A \bar{x}}{\bar{x}^T B \bar{x}} + \lambda \|\bar{x}\|_0 \leq \frac{x^T A x}{x^T B x} + \lambda \|x\|_0 + \eta.
$$

This shows that $\|x\|_0 < \frac{1}{2} + \|\bar{x}\|_0$, which is impossible.

Using Lemma A.2 and noting that $J = \text{supp}(x)$, we derive that
$$
\partial L \Phi(x) \subseteq \left\{ \frac{2A x(x^T B x) - 2B x(x^T A x)}{(x^T B x)^2} + \lambda v + t x : t \in \mathbb{R}, \; v_j = 0 \text{ if } j \in J, \; v_j \in \mathbb{R} \text{ if } j \notin J \right\}.
$$

Denoting $[a]_J = (a_j)_{j \in J} \in \mathbb{R}^{|J|}$, this implies that
$$
\text{dist}(0, \partial L \Phi(x)) \geq \inf_{t \in \mathbb{R}} \left\| \left[ \frac{2A x(x^T B x) - 2B x(x^T A x)}{(x^T B x)^2} \right]_J + t x \right\|.
$$

19
A direct verification shows that
\[ x^\top \left( \frac{2Ax(x^\top Bx) - 2Bx(x^\top Ax)}{(x^\top Bx)^2} \right) = 0, \]
which, together with \( J = \text{supp}(x) \), implies that
\[ (x_j)^\top \left( \frac{2Ax(x^\top Bx) - 2Bx(x^\top Ax)}{(x^\top Bx)^2} \right) = 0. \]

Therefore,
\[ \text{dist}(0, \partial L \Phi(x)) \geq \left\| \frac{2Ax(x^\top Bx) - 2Bx(x^\top Ax)}{(x^\top Bx)^2} \right\|_j = \frac{2}{x^\top Bx} \left\| [Ax]_j - \frac{x^\top Ax}{x^\top Bx} [Bx]_j \right\|. \]

Using \( J = \text{supp}(x) \) again, we have that
\[ [Ax]_j = A_{jj}x_j, \; x^\top Ax = (x_j)^\top A_{jj}x_j, \; [Bx]_j = B_{jj}x_j, \; \text{and } x^\top Bx = (x_j)^\top B_{jj}x_j, \]
and hence
\[ \text{dist}(0, \partial L \Phi(x)) \geq \frac{2}{x^\top Bx} \left\| A_{jj}x_j - \frac{x^\top Ax}{x^\top Bx} B_{jj}x_j \right\|. \]

Now, let
\[ q(z) = z^\top A_{jj} x - \frac{x^\top Ax}{x^\top Bx} z^\top B_{jj} x \text{ for all } z \in \mathbb{R}^{|J|}. \]

Then
\[ \text{dist}(0, \partial L \Phi(x)) \geq \frac{2}{x^\top Bx} \left\| \frac{x^\top Ax}{x^\top Bx} B_{jj}x_j \right\|. \]

where the second inequality follows from the triangle inequality and the last equality holds as \( x, \bar{x} \in \Lambda \) and \( J = \text{supp}(x) = \text{supp} (\bar{x}) \) (and so, \( \|x\|_0 = \|\bar{x}\|_0 \)).

From Lemma 5.5, there exists \( c > 0 \) such that, for all \( z \), \( \|\nabla q(z)\|^2 \geq c q(z) \). Indeed, one can set \( c := \min_{1 \leq j \leq |J|} \left\{ 4|\lambda_j(A_{jj} - \frac{x^\top Ax}{x^\top Bx} B_{jj})| : \lambda_j(A_{jj} - \frac{x^\top Ax}{x^\top Bx} B_{jj}) \neq 0 \right\} \), where \( \lambda_j(Q) \) are the eigenvalues of a matrix \( Q \). Noting that
\[ \frac{q(x_j)}{x^\top Bx} = \frac{(x_j)^\top A_{jj}(x_j) - \frac{x^\top Ax}{x^\top Bx} (x_j)^\top B_{jj} (x_j)}{x^\top Bx} \frac{1}{x^\top Bx} \]
\[ = \frac{x^\top Ax}{x^\top Bx} - \frac{x^\top Ax}{x^\top Bx} \frac{1}{x^\top Bx} \]
\[ = \frac{x^\top Ax}{x^\top Bx} - \frac{x^\top Ax}{x^\top Bx} = \Phi(x) - \Phi(\bar{x}) > 0, \]
one has
\[ \text{dist}(0, \partial L \Phi(x)) \geq \frac{\sqrt{c} q(x_j)^{1/2}}{x^\top Bx} - 2|\Phi(x) - \Phi(\bar{x})|\|B_{jj}x_j\| \]
\[ = [\Phi(x) - \Phi(\bar{x})]^{1/2} \left( \frac{\sqrt{c}}{\sqrt{x^\top Bx}} - 2|\Phi(x) - \Phi(\bar{x})|^{1/2} \|B_{jj}x_j\| \right). \]

Let \( c_1 := \min \{ \sqrt{x^\top Bx} : x \in \Lambda \} \) and \( c_2 := \max \{ \sqrt{x^\top Bx} : x \in \Lambda \} \). By shrinking \( \eta \) if necessary, we can assume that \( \eta \in (0, 1) \) and
\[ 2|\Phi(x) - \Phi(\bar{x})|^{1/2} \|B_{jj}x_j\| \leq 2\eta^{1/2} \|B_{jj}x_j\| \leq \frac{\sqrt{c}}{2c_2}, \]

20
where the first inequality follows by the fact $\Phi(\mathbf{x}) < \Phi(\mathbf{x}) < \Phi(\mathbf{x}) + \eta$. Then, we see that

$$\text{dist}(0, \partial_L \Phi(\mathbf{x})) \geq |\Phi(\mathbf{x}) - \Phi(\mathbf{x})|^{1/2} \left(\frac{\sqrt{c}}{c_2} - \frac{\sqrt{c}}{2c_2}\right) = \frac{\sqrt{c}}{2c_2} |\Phi(\mathbf{x}) - \Phi(\mathbf{x})|^{1/2}.$$ 

Thus, $\Phi$ satisfies the KL property with exponent $\frac{1}{2}$. This shows that, according to Lemma 5.1, $\hat{\Phi}_{GEPS}$ satisfies the KL property with KL exponent $\frac{1}{2}$ at $\mathbf{x}$ for all $\mathbf{x} \in \text{dom} \partial_L \Phi$.

### 5.3. Generalized eigenvalue problem with sparsity constraint

Consider the generalized eigenvalue problem with sparsity constraint

$$\max_{x \in \mathbb{R}^d} \frac{x^\top A_1 x}{x^\top B_1 x} \quad \text{s.t.} \quad \|x\| = 1, \|x\|_0 \leq r,$$

(GEPS)

where $A_1, B_1$ are symmetric matrices such that $A_1$ is positive semidefinite and $B$ is positive definite, and $r > 0$. For this generalized eigenvalue problem with sparsity constraint, the corresponding merit function for the proposed Algorithm 1 takes the form

$$\hat{\Phi}_{GEPS}(x, u) = \frac{x^\top A x}{x^\top B x} + \delta_{C_r}(x) + \delta_{\Lambda}(x) + \rho \|x - u\|^2,$$

where $A = -A_1$ is a symmetric matrix, $B = B_1$ is a positive definite matrix, $\Lambda = \{x \in \mathbb{R}^d : \|x\| = 1\}$, $C_r = \{x \in \mathbb{R}^d : \|x\|_0 \leq r\}$ with $r > 0$, and $\rho \geq 0$. Below, we investigate the KL exponent for this merit function.

**Theorem 5.7.** Consider the function

$$\Phi(x) = \frac{x^\top A x}{x^\top B x} + \delta_{C_r}(x) + \delta_{\Lambda}(x),$$

where $C_r = \{x \in \mathbb{R}^d : \|x\|_0 \leq r\}$, $\Lambda = \{x \in \mathbb{R}^d : \|x\| = 1\}$, and $A, B$ are symmetric matrices with $B$ positive definite. Then $\Phi$ is a KL function with KL exponent $\frac{1}{2}$. In particular, for all $\rho \geq 0$,

$$\hat{\Phi}_{GEPS}(x, u) = \frac{x^\top A x}{x^\top B x} + \delta_{C_r}(x) + \delta_{\Lambda}(x) + \rho \|x - u\|^2$$

satisfies the KL property with KL exponent $\frac{1}{2}$ at $(\mathbf{x}, \mathbf{x})$ for all $\mathbf{x} \in \text{dom} \partial_t \Phi$.

**Proof.** Take any $\mathbf{x} \in \Lambda \cap C_r$. We split the proof into two cases: $\|\mathbf{x}\|_0 = r$ and $\|\mathbf{x}\|_0 < r$.

**Case 1:** $\|\mathbf{x}\|_0 = r$. Let $\delta > 0$ and take any $x \in \Lambda \cap C_r$ with $\|\mathbf{x} - \mathbf{x}\| \leq \delta$. By shrinking $\delta$ if necessary, we have $\text{supp}(\mathbf{x}) \subseteq \text{supp}(\mathbf{x})$. So, $\|\mathbf{x}\|_0 \geq \|\mathbf{x}\|_0 = r$. As $x \in C_r$, we see that $\|\mathbf{x}\|_0 = \|\mathbf{x}\|_0 = r$ and so, $\text{supp}(\mathbf{x}) = \text{supp}(\mathbf{x})$. Then, a similar line of argument as in Theorem 5.6 gives the desired conclusion.

**Case 2:** $\|\mathbf{x}\|_0 < r$. Let $\tilde{I} = \{I \subseteq \{1, \ldots, n\} : \text{supp}(\mathbf{x}) \subseteq I\}$. Clearly, $|\tilde{I}| < +\infty$. Let $\delta > 0$ and take any $x \in \Lambda \cap C_r$ with $\|\mathbf{x} - \mathbf{x}\| \leq \delta$. By shrinking $\delta$ if necessary, we have $\text{supp}(\mathbf{x}) \subseteq \text{supp}(\mathbf{x})$, and so, $J_x := \text{supp}(\mathbf{x}) \in \tilde{I}$. Let $x$ with $\|\mathbf{x} - \mathbf{x}\| < \eta$ and $\Phi(x) < \Phi(x) < \Phi(x) + \eta$. From our choice of $x$, one has $x \in \Lambda$. Moreover, using Lemma A.2, a direct computation gives us that

$$\partial_L \Phi(x) \subseteq \left\{\frac{2A\mathbf{x}(\mathbf{x}^\top B) - 2B\mathbf{x}(\mathbf{x}^\top A\mathbf{x})}{(\mathbf{x}^\top B\mathbf{x})^2} + \lambda v + t \mathbf{x} : t \in \mathbb{R}, \tilde{J} \subseteq \{1, \ldots, n\} \setminus J_x, |\tilde{J}| = r - |J_x|, \quad v_i = 0 \text{ if } i \in J_x \cup \tilde{J}, \text{ and } v_i \in \mathbb{R} \text{ if } i \notin \text{supp}(\mathbf{x}) \cup \tilde{J}\right\}.$$

As $x \in \left(\frac{2A\mathbf{x}(\mathbf{x}^\top B) - 2B\mathbf{x}(\mathbf{x}^\top A\mathbf{x})}{(\mathbf{x}^\top B\mathbf{x})^2}\right)$, $0$,

$$x^\top \left(\frac{2A\mathbf{x}(\mathbf{x}^\top B) - 2B\mathbf{x}(\mathbf{x}^\top A\mathbf{x})}{(\mathbf{x}^\top B\mathbf{x})^2}\right)_{J_x \cup \tilde{J}} = 0.$$
Thus,
\[
\text{dist}(0, \partial_L \Phi(x)) \geq \inf_{t \in \mathbb{R}, J \subseteq \{1, \ldots, n\} \setminus J_x, |J| = r - J_x} \left\| \frac{2Ax(x^T Bx) - 2Bx(x^T Ax)}{(x^T Bx)^2} + tx \right\|_{J_x \cup J} \\
= \inf_{J \subseteq \{1, \ldots, n\} \setminus J_x, |J| = r - J_x} \left\| \frac{2Ax(x^T Bx) - 2Bx(x^T Ax)}{(x^T Bx)^2} \right\|_J.
\]
Using a similar line of argument as in Theorem 5.6, one has
\[
\text{dist}(0, \partial_L \Phi(x)) \geq \inf_{J \supseteq J_x, |J| = r} \left\{ \frac{1}{x^T Bx} \| \nabla q_J(x) \| - \frac{c_J}{x^T Bx} \| B_J x_J \| \right\},
\]
where
\[
q_J(z) = z^T A_J z - \frac{x^T A x}{x^T B x} x^T B_J z + B_J z 
\]
for all \( z \in \mathbb{R}^{|J|} \).
From Lemma 5.5, we see that, for each \( J \supseteq J_x \) with \(|J| = r\), there exists \( c_J > 0 \) such that
\[
\| \nabla q_J(z) \|^2 \geq c_J q_J(z).
\]
Note that \( \{ J : J \supseteq J_x \text{ with } |J| = r \} \subseteq \mathcal{I} := \{ J : J \supseteq J_x \text{ with } |J| = r \} \) (as \( \text{supp}(x) \subseteq \text{supp}(x) \)) and \( |\mathcal{I}| < +\infty \). So, \( c := \min_{J \in \mathcal{I}} c_J > 0 \). Noting from \( \Phi(x) - \Phi(x) > 0 \), for each \( J \supseteq J_x \) with \(|J| = r\), one has
\[
q_J(x_J) = \left( x_J^T A_J x_J - \frac{x^T A x}{x^T B x} x^T B_J x_J \right) \frac{1}{x^T B x} \\
= \left( x^T A x - \frac{x^T A x}{x^T B x} x^T B x \right) \frac{1}{x^T B x} \\
= \frac{x^T A x}{x^T B x} - \frac{x^T A x}{x^T B x} = \Phi(x) - \Phi(x) > 0.
\]
Thus,
\[
\text{dist}(0, \partial_L \Phi(x)) \geq \inf_{J \supseteq J_x, |J| = r} \left\{ \frac{\sqrt{c_J} |x_J|^{1/2}}{x^T B x} - \frac{2|\Phi(x) - \Phi(x)|}{x^T B x} \| B_J x_J \| \right\} \\
= |\Phi(x) - \Phi(x)|^{1/2} \inf_{J \supseteq J_x, |J| = r} \left\{ \frac{c}{\sqrt{x^T B x}} - \frac{2|\Phi(x) - \Phi(x)|^{1/2}}{x^T B x} \| B_J x_J \| \right\}.
\]
Following a similar line of arguments as in Theorem 5.6, we get the desired conclusion.

**Remark 5.8 (Linear convergence of Algorithm 1).** From Proposition 4.9, Theorem 5.6 and Theorem 5.7, one sees that Algorithm 1 exhibits linear convergence when applied to generalized eigenvalue problems with cardinality regularization and generalized eigenvalue problems with sparsity constraints.

### 6. Numerical examples

In the section, we illustrate our proposed method via numerical examples. We first start with an explicit analytic example and use it to demonstrate the behavior of Algorithm 1 as well as the effect of the inertial parameters. Then, we examine the performance of the algorithm for the sparse eigenvalue optimization model. All the numerical tests are conducted on a computer with a 2.8 GHz Intel Core i7 and 8 GB RAM, equipped with MATLAB R2015a.

#### 6.1. Analytic examples

**Example 6.1 (An illustrative nonsmooth optimization example with multiple stationary points).** Consider
\[
\max_{x \in \mathbb{R}^m} \sum_{i=1}^m |x_i - 1| + \gamma \sum_{i=1}^m \frac{x_i + 1}{x_i^2 + 2x_i + 5} \quad \text{s.t.} \quad 0 \leq x \leq 50,
\]
\[
(EP)
\]
where $\gamma > 0$. Note that $x_i^2 + 2x_i + 5 = (x_i + 1)^2 + 2^2 \geq 4(x_i + 1)$, with equality for $x_i = 1$, for all $i \in \{1, \ldots, m\}$. Direct verification shows that $\mathbf{x} = (1, \ldots, 1)$ is the global solution of this problem. This problem has multiple (lifted) stationary points. As an illustration, Figure 1 depicts the graph of this function with $m = 2$ and $\gamma = 10$ in $[0,5] \times [0,5]$. From the graph, one sees that $(0,0)$ is a local minimizer and $(1,1)$ is the global maximizer. Moreover, direct computation shows that, in the case $m = 2$, this problem has four (lifted) stationary points on $\mathbb{R}^2$, which are $(0,0), (0,1), (1,0)$ and $(1,1)$. We also note that the three (lifted) stationary points which are not global solutions, $(0,0), (0,1)$ and $(1,0)$ do not satisfy the strong stationary condition (4.8).

The equivalent non-fractional problem is given by

$$
\max_{0 \leq x \leq 50, y \in \mathbb{R}^m} - \sum_{i=1}^{m} |x_i - 1| + \sum_{i=1}^{m} \gamma \left( 2y_i \sqrt{x_i + 1} - y_i^2 (x_i^2 + 2x_i + 5) \right).
$$

Note that this example satisfies Assumption A with $\alpha_i = 1/2$ and $\beta_i = 2$, for all $i \in \{1, \ldots, m\}$. Let $\gamma = 10$, $x_0 = x_{-1}$, $\delta = 1$, $\tau_n = 25$, $\nu_n = 0$, $n \in \mathbb{N}$. Let $f_i(x_i) = 10(x_i + 1)$ and $g_i(x_i) = x_i^2 + 2x_i + 5$, for all $i \in \{1, \ldots, m\}$. Clearly, $M_i = \max_{0 \leq x_i \leq 50} f_i(x_i) = 510$ and $m_i = \min_{0 \leq x_i \leq 50} g_i(x_i) = 5$, for all $i \in \{1, \ldots, m\}$. Thus, one can verify that $\tau_n \equiv 25 \geq \delta + \max_{1 \leq i \leq m} \{ \alpha_i \frac{M_i}{m_i} + \frac{1}{2} \beta_i \frac{M_i}{m_i} \}$, and $z_{i,n} = x_{i,n} + \nu_n (x_{i,n} - x_{i,n-1}) = x_{i,n}$, for all $i \in \{1, \ldots, m\}$. Note that

$$
p(s) := \arg\min_{0 \leq x \leq 50} \{ |x - 1| + (x - s)^2 \} = \begin{cases} 50 & \text{if } s \geq \frac{101}{2}, \\ s - \frac{1}{2} & \text{if } \frac{3}{2} < s < \frac{101}{2}, \\ 1 & \text{if } \frac{1}{2} < s \leq \frac{3}{2}, \\ s + \frac{1}{2} & \text{if } -\frac{1}{2} < s \leq \frac{1}{2}, \\ 0 & \text{if } s < -\frac{1}{2}. \end{cases}
$$

Then, Algorithm 1 can be simplified to the following compact form: set $x_{n+1} = (x_{1,n+1}, \ldots, x_{m,n+1})$ with

$$
x_{i,n+1} = \arg\max_{x \geq 0} \left\{ -|x - 1| - \left[ x - (z_{i,n} + \gamma \frac{(-x_{i,n}^2 - 2x_{i,n} + 3)}{2\tau_n (x_{i,n}^2 + 2x_{i,n} + 5)^2}) \right]^2 \right\} = p \left( x_{i,n} + \frac{(-x_{i,n}^2 - 2x_{i,n} + 3)}{5(x_{i,n}^2 + 2x_{i,n} + 5)^2} \right).
$$

We randomly generate initial points in $[0,50]^m$ and perform Algorithm 1. For all the initial points, the algorithm produces a sequence $(x_n)_{n \geq 0}$ converging to the true global maximizer $\mathbf{x} = (1, \ldots, 1)$. Figure 2 depicts the convergence behavior for the case $m = 2$ and $\gamma = 10$, with initial points $(50,50), (20,20), (0,1), (1,0)$ and $(0,0)$ by plotting out the Euclidean distance to the true solution $(1,1)$ per iteration. Interestingly, the algorithm converges to the true global maximizer even when the initial points are selected as $(0,0), (0,1)$ and $(1,0)$, which are already stationary points. We now illustrate the behavior of Algorithm 1 by varying the inertial parameters. To do this, we fix $m = 2$ and $\gamma = 10$ and an $\alpha \in (0,1)$. We set $\nu_n \equiv \alpha_n^2 = \frac{\gamma}{2}$. Starting with the initialization $x_0 = (50,50)$, we then run Algorithm 1 with different values for $\alpha \in [0,1]$. Figure 3 depicts the distance, in the log scale, between the sequence of iterates $(x_n)_{n \geq 0}$ and the solution $\mathbf{x} = (1,1)$, for $\alpha \in \{0,0.5,0.7,0.99\}$. As one can see from the figure, as $\alpha$ increases and approaches 1, the algorithm tends to converge faster. 

\[^{3}\text{This is due to the fact that } (1,1) \text{ is the only (lifted) stationary point which satisfies the stronger stationary condition (4.8).}\]
6.2. Sparse generalized eigenvalue problems

As another illustration of our algorithm, following [24], we consider a sparse generalized eigenvalue problem that arises from binary classification using sparse Fisher discriminant analysis. Consider $p$ observations $z_1, \ldots, z_p$ with $z_i \in \mathbb{R}^d$, $i \in \{1, \ldots, p\}$, each of which belongs to one of two distinct classes. Let $I_k \subseteq \{1, \ldots, p\}$ contain the indices of the observations in class $k$, with $p_k = |I_k|$, $k = 1, 2$, and $p_1 + p_2 = p$. Let $\mu_k = \frac{1}{p_k} \sum_{i \in I_k} z_i$, for $k = 1, 2$. The so-called within-class and between-class covariance matrices are given by

$$V_w = \frac{1}{p} \sum_{k=1}^{2} \sum_{i \in I_k} (z_i - \mu_k)(z_i - \mu_k)^\top$$
$$V_b = \frac{1}{p} \sum_{k=1}^{2} p_k \mu_k \mu_k^\top.$$

The classification problem using sparse Fisher discriminant analysis (SFDA) then seeks a low dimensional projection of the observations such that the between-class variance is large relative to the within-class variance. Mathematically, it solves

$$\max_{x \in \mathbb{R}^d} \frac{x^\top V_b x}{x^\top V_w x} - \lambda \phi(x) \quad \text{s.t.} \quad \|x\| = 1,$$

(SFDA)

where $\phi$ is a regularization function inducing sparsity, and $\lambda > 0$. This is a sparse generalized eigenvalue problem with $A = V_b$ and $B = V_w$. Here, we consider two specific sparse regularization functions: $\phi(x) = \|x\|_0$, and $\phi(x) = \delta_{C_r}(x)$ with $C_r = \{x \in \mathbb{R}^d : \|x\|_0 \leq r\}$ and $r > 0$.

In the case where $\phi(x) = \delta_{C_r}(x)$, [24] proposed a truncated Rayleigh flow method (TRFM) for solving the above sparse generalized eigenvalue problem and showed that this method converges when the initial point $x_0$ is close.
enough to a global solution of the underlying nonconvex optimization problem. Moreover, they also showed that, in this case, the method exhibits R-linear convergence. We note that, in general, it is hard to theoretically guarantee whether an initial point $x_0$ is chosen to be close enough to a global solution, in order to ensure the convergence of the algorithm. On the other hand, Algorithm 1 can be applied to (SFDA) with both $\phi(x) = \|x\|_0$ and $\phi(x) = \delta_{C_v}(x)$, and Remark 5.8 shows that Algorithm 1 converges linearly regardless of the choice of the initial points.

### 6.2.1. Sparsity constrained case

In this subsection, we consider the generalized eigenvalue problem with sparsity constraints, that is, (SFDA) with $\phi(x) = \delta_{C_v}(x)$. In this setting, Algorithm 1 reads as

$$\text{set } x_{n+1} = P_{\Lambda \cap C_v} \left( z_n + \frac{1}{\tau_n} \frac{x_n^T V_i x_n}{(x_n^T V_i x_n)^2} \left[ x_n^T V_i x_n V_i x_n - V_i x_n \right] \right) \text{ with } z_n = x_n + \nu_n (x_n - x_{n-1}),$$

where $P_{\Lambda \cap C_v}(a)$ is the Euclidean projection of $a$ on $\Lambda \cap C_v$ and it has a closed form solution. Indeed, denoting $T_r(a)$ as the Euclidean projection of $a$ to $C_v$, it is known that, for $a = (a_1, \ldots, a_d) \in \mathbb{R}^d$, $(T_r(a))_i = a_i$ for the $r$ largest components in absolute value of $a$, and $(T_r(a))_i = 0$ otherwise. Then

$$P_{\Lambda \cap C_v}(a) = \begin{cases} \left\{ \frac{a}{\|a\|} : a \in T_r(a) \right\} & \text{if } a \neq 0, \\ \Lambda \cap C_v & \text{if } a = 0. \end{cases}$$

This can be seen, for example, by noting that $P_{\Lambda \cap C_v}(a) = \arg\min \left\{ \frac{1}{2}\|x - a\|^2 : x \in \Lambda \cap C_v \right\} = \arg\min \left\{ \langle a, x \rangle : x \in \Lambda \cap C_v \right\}$, and applying [18, Proposition 13].

In our simulation, we adopt the same setting as in [24]: we set $\mu_1 = 0$, $\mu_2 = (\mu_{2,1}, \ldots, \mu_{2,d})^T$ with $\mu_{2,j} = 0.5$ for $j \in \{2, 4, \ldots, 40\}$ and $\mu_{2,j} = 0$ otherwise. Let $\Sigma$ be a block diagonal covariance matrix with five blocks, each of dimension $(d/5 \times d/5)$. The $(j, j')$-th element of each block takes the value $0.8^{|j-j'|}$. As explained in [24], this covariance structure is intended to mimic the covariance structure of a gene expression data. The observation data are simulated as $z_i \sim N(\mu_i, \Sigma)$ for $i \in I_k, k = 1, 2$.

We use our proposed inertial proximal method (Algorithm 1) and the truncated Rayleigh flow method (TRFM) for solving (SFDA) with $\phi(x) = \delta_{C_v}(x)$, where we set $r = 50$, $p_1 = p_2 = 500$, $p = p_1 + p_2 = 1000$, and $d = 2000$.

- For Algorithm 1, we use the initial point $x_0 = (1/\sqrt{r}, \ldots, 1/\sqrt{r}, 0, \ldots, 0) \in \mathbb{R}^d$. Direct verification shows that Assumption A is satisfied with $\alpha_1 = 0$ and $\beta_1 = 2\lambda_{\max}(V_w)$, $\tau_n = 1 + \frac{\|V_i x_n\|_2}{(x_n^T V_i x_n)^2} \lambda_{\max}(V_w)$, $\nu = 0.4999 < \frac{\tau}{\tau_n}$ and $\nu_n = \frac{\nu}{\tau_n}$. We stop the algorithm when either the iterations reach the maximum iteration number $6000$ or the quantity $\|x_{n+1} - x_n\|$ is less than $10^{-6}$.

- For (TRFM), we use the same initial point $x_0$ as in Algorithm 1. We also use the same termination criteria as in Algorithm 1.

We run TRFM and Algorithm 1 for 50 trials. Table 1 summarizes the output of the two methods by listing the average value for

(i) the sparsity level of the solution (round to the nearest integer): the number of entries of the computed solution with absolute value larger than $10^{-6}$;

(ii) the objective value of the computed solution;

(iii) the CPU time measured in seconds;

(iv) the number of iterations used (round to the nearest integer).

From Table 1, one can see that Algorithm 1 is competitive with the TRFM method and produces a solution with better quality in terms of sparsity and final objective value (note that (SFDA) is a maximization problem). Moreover, Algorithm 1 also uses less CPU time and number of iterations. As an illustration, we also plot $\|x_n - x^*\|$ against the number of iterations $n$, in logarithmic scale, where $x^*$ is the approximated solution produced by the corresponding algorithm. Figure 4 supports the theoretical finding that Algorithm 1 exhibits linear convergence in this case.
6.2.2. Sparse generalized eigenvalue problem with cardinality regularization

In this subsection, we consider the generalized eigenvalue problem with cardinality regularization, that is, (SFDA) with $\phi(x) = \|x\|_0$. In this setting, Algorithm 1 reads

$$x_{n+1} = \text{argmax} \left\{ -\lambda \|x\|_0 - \tau_n \|x - z_n - \frac{1}{2\tau_n} w_n\|^2 : \|x\| = 1 \right\}$$

$$= \text{argmax} \left\{ -\lambda \|x\|_0 + \langle 2\tau_n z_n + w_n, x \rangle : \|x\| = 1 \right\}$$

with $\lambda > 0$, $z_n = x_n + \nu_n (x_n - x_{n-1})$ and

$$w_n = \frac{x_n^\top V_b x_n}{(x_n^\top V_b x_n)^2} \left[ x_n^\top V_w x_n V_b x_n - V_w x_n \right].$$

We note that, for each $a \in \mathbb{R}^d$, the optimization problem $\text{argmax} \{-\lambda \|x\|_0 + \langle a, x \rangle : \|x\| = 1\}$ has a closed form solution [23, Proposition 6]. In our numerical experiment, we set $\lambda = 3.5e - 2$. We also generate the data as in the previous subsection, using the same initial point, parameters $\tau_n$, $\nu_n$ and $\delta$, and employing the same termination criteria.

We run Algorithm 1 for 50 trials. Table 2 summarizes the output of the method where the meanings of the items are the same as in the previous subsection.

| Algorithm 1 | Sparsity level of the computed solution | Objective value of the computed solution | CPU time (sec) | Number of iterations |
|-------------|---------------------------------------|----------------------------------------|----------------|---------------------|
| TRFM        | 26                                    | 11.5051                                | 6.1950         | 1202                |
| Algorithm 1 | 23                                    | 12.5158                                | 3.9684         | 564                 |

Table 2: Computation results for (SFDA) with cardinality regularization

We also plot out Euclidean distance between $x_n$ and $x^*$ per iteration in log scale, which supports the theoretical finding that Algorithm 1 exhibits linear convergence for this problem.
Figure 5: Euclidean distance between $x_n$ and $x^*$ in every iteration

Acknowledgment: The authors are grateful for Dr. Qia Li for the help and discussions for the applications of sparse Fisher discriminant analysis.

A. Appendix

In this appendix, we collect several useful calculus rules for limiting subdifferentials which are used in this paper.

Lemma A.1 (Calculus rules). Let $f, g: \mathcal{H} \to (-\infty, +\infty]$ be proper lower semicontinuous functions and let $x \in \mathcal{H}$. Then the following statements hold:

(i) (Separable sum rule) If $f(x) = \sum_{i=1}^{m} f_i(x_i)$ with $x = (x_1, \ldots, x_m)$, then $\partial_L f(x) = \partial_L f_1(x_1) \times \cdots \times \partial_L f_m(x_m)$.

(ii) (Sum rule) If $f$ is finite at $x$ an $g$ is Lipschitz continuous around $x$, then $\partial_L (f + g)(x) \subseteq \partial_L f(x) + \partial_L g(x)$.

(iii) (Sign rule) If $f$ is Lipschitz continuous around $x$ and $\partial f$ is nonempty-valued around $x$, then $\partial_L (-f)(x) \subseteq -\partial_L f(x)$.

(iv) (Quotient rule) Suppose that $f$ and $g$ are Lipschitz continuous around $x$, and $g(x) \neq 0$. If $\partial f$ is nonempty-valued around $x$, then

$$
\partial_L \left( \frac{-f}{g} \right)(x) \subseteq -\frac{g(x)\partial_L f(x) + \partial_L (f(x)g(x))}{g(x)^2}.
$$

If $f$ is strictly differentiable at $x$, then

$$
\partial_L \left( \frac{-f}{g} \right)(x) = -\frac{g(x)\nabla f(x) + \partial_L (f(x)g(x))}{g(x)^2}.
$$

(v) (Chain rule and square root rule) If $f$ is Lipschitz continuous around $x$ and $\theta: \mathbb{R} \to \mathbb{R}$ is continuously differentiable around $f(x)$, then $\partial_L (\theta \circ f)(x) = \partial_L (\theta'(f(x)) f)(x)$. In particular, if $f$ is Lipschitz continuous around $x$ and $f(x) > 0$, then

$$
\partial_L \left( \sqrt{f} \right)(x) = \frac{\partial_L f(x)}{2\sqrt{f(x)}} \quad \text{and} \quad \partial_L \left( -\sqrt{f} \right)(x) = \frac{\partial_L (-f)(x)}{2\sqrt{f(x)}}.
$$

Proof. (i): This is given in [21, Proposition 10.5].

(ii): This follows from [19, Proposition 1.107(ii) and Theorem 2.33].

(iii): This is an application of [20, Corollary 3.4] with $\varphi_1 \equiv 0$ and $\varphi_2 = f$. 

27
(iv): We first have from [19, Proposition 1.111(ii)] and (ii) that
\[
\partial_L \left( \frac{-f}{g} \right)(x) = \frac{\partial_L (-g(x)f + f(x)g)(x)}{g(x)^2} \subseteq \frac{\partial_L (-g(x)f)(x) + \partial_L (f(x)g)(x)}{g(x)^2}.
\] (A.1)
Assume that \(\hat{f}\) is nonempty-valued around \(x\). Then, if \(g(x) > 0\), \(\partial_L (-g(x)f)(x) = g(x) \partial_L (-f)(x) \subseteq -g(x) \partial_L f(x)\) due to (iii). If \(g(x) \leq 0\), then \(-g(x) \geq 0\) and \(\partial_L (-g(x)f)(x) = -g(x) \partial_L f(x)\). Thus, we obtain the desired inclusion. The remaining conclusion follows from the equality in (A.1) and (ii).

(v): The chain rule is given in [19]. The two square root rules follow by letting \(\theta(t) = \sqrt{t}\) and \(\theta(t) = -\sqrt{t}\), respectively.

Lemma A.2. Let \(\Lambda := \{x \in \mathbb{R}^d : \|x\| = 1\}\) and \(C_r := \{x \in \mathbb{R}^d : \|x\|_0 \leq r\}\). Then the following statements hold:

(i) \(\forall x \in \mathbb{R}^d \partial_L^\infty(\| \cdot \|_0)(x) = \partial_L(\| \cdot \|_0)(x) = \{v : v_j = 0 \text{ if } j \in \text{supp}(x)\}\) with \(\text{supp}(x) = \{j : x_j \neq 0\}\).

(ii) \(\forall x \in \Lambda \partial_L^\infty \delta_{\Lambda}(x) = \partial_L \delta_{\Lambda}(x) = \{x : t \in \mathbb{R}\}\).

(iii) If \(\|x\|_0 = r\), then
\[
\partial_L^\infty \delta_{C_r}(x) = \partial_L \delta_{C_r}(x) = \{v : v_j = 0 \text{ if } j \in \text{supp}(x)\}.
\]
If \(\|x\|_0 < r\), then
\[
\partial_L^\infty \delta_{C_r}(x) = \partial_L \delta_{C_r}(x) = \{v : \exists \tilde{J} \subseteq \{1, \ldots, d\} \setminus \text{supp}(x) \text{ with } |\tilde{J}| = r - \|x\|_0, v_j = 0 \text{ if } j \in (\text{supp}(x) \cup \tilde{J})\}.
\]
(iv) \(\forall x \in \Lambda \partial_L(\| \cdot \|_0 + \delta_{\Lambda})(x) \subseteq \partial_L(\| \cdot \|_0)(x) + \partial_L \delta_{\Lambda}(x)\).

(v) \(\forall x \in \Lambda \cap C_r \partial_L(\delta_{C_r} + \delta_{\Lambda})(x) \subseteq \partial_L \delta_{C_r}(x) + \partial_L \delta_{\Lambda}(x)\).

Proof. (i): The limiting subdifferential formula for \(\| \cdot \|_0\) can be found in [14, Section 3]. The formula for the horizon subdifferential can be verified directly.

(ii): Follows by direct verification.

(iii): The limiting subdifferential formula for \(\delta_{C_r}\) can be found in [3, Theorem 3.9]. The formula for the horizon subdifferential can be verified directly.

(iv)\&(v): We deduce from (i), (ii), and (iii) that, for all \(x \in \Lambda\), \((-\partial_L^\infty(\| \cdot \|_0)(x)) \cap \partial_L^\infty \delta_{\Lambda}(x) = \{0\}\) and that, for all \(x \in \Lambda \cap C_r\), \((-\partial_L^\infty \delta_{C_r}(x)) \cap \partial_L^\infty \delta_{\Lambda}(x) = \{0\}\). The conclusion then follows from [21, Corollary 10.9].

References

[1] H. Attouch and J. Bolte, On the convergence of the proximal algorithm for nonsmooth functions involving analytic features, Math. Program. Ser. B 116(1–2), 5–16 (2009).
[2] H. Attouch, J. Bolte, and B.F. Svaiter, Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods, Math. Program. Ser. A 137(1–2), 91–129 (2013).
[3] H.H. Bauschke, D.R. Luke, H.M. Phan, and X. Wang, Restricted normal cones and sparsity optimization with affine constraints, Found. Comput. Math. 14(1), 63–83 (2014).
[4] A. Beck and N. Hallak, Proximal mapping for symmetric penalty and sparsity, SIAM J. Optim. 28(1), 496–527 (2018).
[5] H.P. Benson, Solving sum of ratios fractional programs via concave minimization, J. Optim. Theory Appl. 135(1), 1–17 (2007).
[6] J. Bolte, S. Sabach, and M. Teboulle, Proximal alternating linearized minimization for nonconvex and nonsmooth problems, Math. Program. Ser. A 146(1–2), 459–494 (2014).
[7] R.I. Boţ and E.R. Csetnek, Proximal-gradient algorithms for fractional programming, Optimization 66(8), 1383–1396 (2017).
[8] R.I. Boţ, M.N. Dao, and G. Li, Extrapolated proximal subgradient algorithms for nonconvex and nonsmooth fractional programs, arXiv:2003.04124 (2020).
[9] J.-P. Crouzeix, J.A. Ferland, and S. Schaible, An algorithm for generalized fractional programs, *J. Optim. Theory Appl.* **47**(1), 35–49 (1985).

[10] D. D’Acunto and K. Kurdyka, Explicit bounds for the Łojasiewicz exponent in the gradient inequality for polynomials, *Ann. Pol. Math.* **87**, 51–61 (2005).

[11] W. Dinkelbach, On nonlinear fractional programming, *Management Science* **13**, 492–498 (1967).

[12] T. Ibaraki, Parametric approaches to fractional programs, *Math. Prog.* **26**(3), 345–362 (1983).

[13] K. Kurdyka, On gradients of functions definable in o-minimal structures, *Annales de l’institut Fourier* (Grenoble) **48**(3), 769–783 (1998).

[14] H.Y. Le, Generalized subdifferentials of the rank function, *Optim. Lett.* **7**(4), 731–743 (2013).

[15] S. Łojasiewicz, Une propriété topologique des sous-ensembles analytiques réels, *Les Équations aux Dérivées Partielles*, Éditions du Centre National de la Recherche Scientifique, Paris, 87–89 (1963).

[16] G. Li and T.K. Pong, Calculus of the exponent of Kurdyka–Łojasiewicz inequality and its applications to linear convergence of first-order methods, *Found. Comput. Math.* **18**(5), 1199–1232 (2018).

[17] M. Locatelli, Alternative branching rules for some nonconvex problems, *Optim. Methods Softw.* **30**(2), 365–378 (2015).

[18] R. Luss and M. Teboulle, Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint, *SIAM Rev.* **55**(1), 65–98 (2013).

[19] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation I. Basic Theory*, Springer, Berlin (2006).

[20] B.S. Mordukhovich, N.M. Nam, and N.D. Yen, Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming, *Optimization* **55**(5–6), 685–708 (2006).

[21] R.T. Rockafellar and R.J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin (1998).

[22] S. Schaible, Fractional programming II. On Dinkelbach’s algorithm, *Management Sci.* **22**(8), 868–873 (1975/76).

[23] J. Song, P. Babu, and D.P. Palomar, Sparse generalized eigenvalue problem via smooth optimization, *IEEE Trans. Signal Proc.* **63**(7), 1627–1642 (2015).

[24] K.M. Tan, Z. Wang, H. Liu, and T. Zhang, Sparse generalized eigenvalue problem: Optimal statistical rates via truncated Rayleigh flow, *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **80**(5), 1057–1086 (2018).

[25] A. Zappone and E. Jorswieck, Energy efficiency in wireless networks via fractional programming theory, *Found. Trends Commun. Inf. Theory* **11**(3–4), 185–396 (2014).

[26] L.H. Zhang, On optimizing the sum of the Rayleigh quotient and the generalized Rayleigh quotient on the unit sphere, *Comput. Optim. Appl.* **54**(1), 111–139 (2013).