Fast Algorithm for Constrained Linear Inverse Problems

Sheriff, Mohammed Rayyan; Redel, Floor Fenne; Esfahani, Peyman Mohajerin

Publication date
2025

Document Version
Final published version

Published in
Journal of Machine Learning Research

Citation (APA)
Sheriff, M. R., Redel, F. F., & Esfahani, P. M. (2025). Fast Algorithm for Constrained Linear Inverse Problems. *Journal of Machine Learning Research*, *26*, 1-41.

Journal of Machine Learning Research 26 (2025) 1-41 Submitted 12/22; Revised 6/25; Published 6/25

Fast Algorithm for Constrained Linear Inverse Problems

Mohammed Rayyan Sheriff M.R.SHERIFF@TUDELFT.NL
Floor Fenne Redel F.F.REDEL@STUDENT.TUDELFT.NL
Peyman Mohajerin Esfahani P.MOHAJERINESFAHANI@TUDELFT.NL
Delft Center for Systems & Control
Delft University of Technology
Delft, The Netherlands

Editor: Pradeep Ravikumar

Abstract

We consider the constrained Linear Inverse Problem (LIP), where a certain atomic norm (like the ℓ_1 norm) is minimized subject to a quadratic constraint. Typically, such cost functions are non-differentiable which makes them not amenable to the fast optimization methods existing in practice. We propose two equivalent reformulations of the constrained LIP with improved convex regularity: (i) a smooth convex minimization problem, and (ii) a strongly convex min-max problem. These problems could be solved by applying existing acceleration-based convex optimization methods which provide better $O\left(1/k^2\right)$ theoretical convergence guarantee, improving upon the current best rate of $O\left(1/k\right)$. We also provide a novel algorithm named the Fast Linear Inverse Problem Solver (FLIPS), which is tailored to maximally exploit the structure of the reformulations. We demonstrate the performance of FLIPS on the classical problems of Binary Selection, Compressed Sensing, and Image Denoising. We also provide open source `MATLAB` and `PYTHON` package for these three examples, which can be easily adapted to other LIPs.

Keywords: linear inverse problems, min-max problems, sparse coding, image processing.

1. Introduction

Linear Inverse Problems simply refer to the task of recovering a signal from its noisy linear measurements. LIPs arise in many applications, such as image processing (Elad and Aharon, 2006; Yaghoobi and Davies, 2009; Aharon et al., 2006; Olshausen and Fieldt, 1997), compressed sensing (Donoho, 2006a; Candès and Tao, 2006; Candès and Wakin, 2008; Gleichman and Eldar, 2011), recommender systems (Recht et al., 2010), and control system engineering (Nagahara et al., 2015). Formally, given a signal $f \in \mathbb{H}$, and its noisy linear measurements $\mathbb{R}^n \ni x = \phi(f) + \xi$, where, $\phi : \mathbb{H} \longrightarrow \mathbb{R}^n$ is a linear measurement operator and $\xi \in \mathbb{R}^n$ is the measurement noise. The objective is to recover the signal f given its noisy measurements x, and the measurement operator ϕ. Of specific interest is the case when the number of measurements available are fewer than the ambient dimension of the signal, i.e., $n < \dim(\mathbb{H})$. In which case, we refer to the corresponding LIP as being 'ill-posed' since

there could be potentially infinitely many solutions satisfying the measurements even for the noiseless case. In principle, one cannot recover a generic signal f from its measurements if the problem is ill-posed. However, the natural signals we encounter in practice often have much more structure to be exploited. For instance, natural images and audio signals tend to have a sparse representation in a well-chosen basis, matrix valued signals encountered in practice have low rank, etc. Enforcing such a low-dimensional structure into the recovery problem often suffices to overcome its ill-posedness. This is done by solving an optimization problem with an objective function that promotes the expected low-dimensional structure in the solution like sparsity, low-rank, etc. It is now well established that under very mild conditions, such optimization problems and even their convex relaxations often recover the true signal almost accurately (Donoho, 2006b,a; Candès and Wakin, 2008).

1.1 Problem setup

Given $x \in \mathbb{R}^n$, the linear operator $\phi : \mathbb{H} \longrightarrow \mathbb{R}^n$, and $\epsilon > 0$, the object of interest in this article is the following optimization problem

$$\begin{cases} \underset{f \in \mathbb{H}}{\text{argmin}} & c(f) \\ \text{subject to} & \|x - \phi(f)\| \leqslant \epsilon, \end{cases} \tag{1}$$

where $\mathbb{H}$ is some finite-dimensional Hilbert space with the associated inner-product $\langle \cdot\ ,\ \cdot \rangle$. The constraint $\|x - \phi(f)\| \leqslant \epsilon$ is measured using the norm derived from an inner product on $\mathbb{R}^n$ (it is independent from the inner product $\langle \cdot\ ,\ \cdot \rangle$ on the Hilbert space $\mathbb{H}$).

The objective function is the mapping $c : \mathbb{H} \longrightarrow \mathbb{R}$ which is known to promote the low-dimensional characteristics desired in the solution. For example, if the task is to recover a sparse signal, we chose $c(\cdot) = \|\cdot\|_1$; if $\mathbb{H}$ is the space of matrices of a fixed order, then $c(\cdot) = \|\cdot\|_*$ (the Nuclear-norm) if low-rank matrices are desired (Candès and Recht, 2009). In general, the objective function is assumed to be *norm like*.

Main challenges of (1) **and existing state of the art methods to solve it**. One of the main challenges to tackle while solving (1) is that, the most common choices of the cost function $c(\cdot)$ like the ℓ_1 norm, are not differentiable everywhere. In particular, the issue of non-differentiability gets amplified since it is prevalent precisely at the suspected optimal solution (sparse vectors). Thus, canonical gradient-based schemes do not apply to (1) with such cost functions. A common workaround is to use the notion of sub-gradients instead, along with a diminishing step-size. However, the Sub-Gradient Descent method (SGD) for generic convex problems converges only at a rate of $O(1/\sqrt{k})$ (Boyd et al., 2003). For high-dimensional signals like images, this can be tiringly slow since the computational complexity scales exponentially with the signal dimensions.

1.2 Existing methods

The current best algorithms overcome the bottleneck of non differentiability in (1) by instead working with the more flexible notion of *proximal gradients*, and applying them to a suitable reformulation of the LIP (1). We primarily focus on two state-of-the-art methods in this

article: the Chambolle-Pock algorithm (CP) (Chambolle and Pock, 2010) and the C-SALSA (Afonso et al., 2009) algorithm, that solve the LIP (1).

(i) *The Chambolle-Pock algorithm*: Using the *convex indicator function* $\mathbb{1}_{B[x,\epsilon]}(\cdot)$ of $B[x,\epsilon]$,[1] the constraints in LIP (1) are incorporated into the objective function to get

$$\min_{f\in\mathbb{H}} \quad c(f) + \mathbb{1}_{B[x,\epsilon]}(\phi(f)). \tag{2}$$

The indicator function $\mathbb{1}_S(\cdot)$ of any closed convex set S is proper and lower-semicontinuous. Consequently, its convex-conjugate satisfies $(\mathbb{1}_S^*)^* = \mathbb{1}_S$. Since $\mathbb{1}^*_{B[x,\epsilon]}(u) = \langle x, u\rangle + \epsilon\|u\|$, for $B[x,\epsilon]$ in particular, we have

$$\mathbb{1}_{B[x,\epsilon]}(\phi(f)) \;=\; \max_{u\in\mathbb{R}^n} \quad \langle \phi(f), u\rangle - \big(\langle x, u\rangle + \epsilon\|u\|\big). \tag{3}$$

Incorporating (3) in (2), the LIP reduces to the following equivalent min-max formulation

$$\min_{f\in\mathbb{H}} \max_{u\in\mathbb{R}^n} \quad c(f) + \langle \phi(f), u\rangle - (\langle x, u\rangle + \epsilon\|u\|). \tag{4}$$

The min-max problem (4) falls under a special subclass of convex-concave min-max problems with bi-linear coupling between the minimizing (f) and maximizing (u) variables. A primal-dual algorithm was proposed in (Chambolle and Pock, 2010, 2016) to solve such problems under the condition that the mappings $f \longmapsto c(f)$ and $u \longmapsto \langle x\ ,\ u\rangle + \epsilon\|u\|$ are proximal friendly. It turns out that in many relevant problems particularly where $c(f) = \|f\|_1$ the proximal operator of $f \longmapsto c(f)$ is indeed easily computable (Beck and Teboulle, 2009). Moreover, the proximal operator for the mapping $u \longmapsto \langle x\ ,\ u\rangle + \epsilon\|u\|$ corresponds to block soft-thresholding, and is also easy to implement. Under such a setting the CP algorithm has an ergodic convergence rate of $O(1/k)$ for the duality gap of (4). This is already an improvement over the $O(1/\sqrt{k})$ rate in canonical sub-gradient "descent" algorithms, and is currently the best convergence guarantee that exists for Problem (1).

(ii) The C-SALSA algorithm: The Constrained Split Augmented Lagrangian Shrinkage Algorithm (C-SALSA) (Afonso et al., 2009) is an algorithm in which the Alternating Direction Method of Multipliers (ADMM) is applied to problem (2). For this problem, ADMM solves the LIP based on variable splitting using an Augmented Lagrangian Method (ALM). In a nutshell, the algorithm iterates between optimizing the variable f and the Lagrange multipliers until they converge. Even though the convergence rates for C-SALSA are not better than that of the CP algorithm, it is empirically found to be fast. Of particular interest is the case when ϕ satisfies $\phi^\top\phi = \mathbb{I}_n$, in which case, further simplifications in the algorithm can be done that improve its speed for all practical purposes.

One of the objectives of this work is to provide an algorithm that is demonstrably faster than the existing methods, particularly for large scale problems. Given that solving an LIP is such a commonly arising problem in signal processing and machine learning, a fast and easy to implement is always desirable. This article precisely caters to this challenge. In (Sheriff and Chatterjee, 2020), the LIP (1) was equivalently reformulated as a convex-concave min-max problem:

$$\min_{h\in B_c} \sup_{\lambda\in\Lambda} \quad 2\sqrt{\langle \lambda\ ,\ x\rangle - \epsilon\,\|\lambda\|} \;-\; \langle \lambda\ ,\ \phi(h)\rangle\,, \tag{5}$$

1. The convex indicator function $\mathbb{1}_S(z)$ of a given convex set S is $\mathbb{1}_S(z) = 0$ if $z \in S$; $= +\infty$ if $z \notin S$.

where $B_c = \{h \in \mathbb{H} : c(h) \leqslant 1\}$ and $\Lambda := \{\lambda \in \mathbb{R}^n : \langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\| > 0\}$. A solution to the LIP (1) can be computed from a saddle point (h^*, λ^*) of the min-max problem (5). Even though primal-dual schemes like Gradient Descent-Ascent with appropriate step-size can be used to compute a saddle-point of the min-max problem (5); such generic methods fail to exploit the specific structure of the min-max form (5). It turns out that equivalent problems with better convex regularity (like smoothness) can be derived from (5) by exploiting the specific nature of this min-max problem.

1.3 Contribution

In view of the existing methods mentioned above, we summarize the contributions of this work as follows:

(a) **Exact reformulations with improved $O(1/k^2)$ convergence rates.** We build upon the min-max reformulation (5) and proceed further on two fronts to obtain equivalent reformulations of the LIP (1) with better convex regularities. These reformulations open the possibility for applying acceleration based methods to solve the LIP (1) with faster rates of convergence $O(1/k^2)$, which improves upon the existing best rate of $O(1/k)$.
   (i) *Exact smooth reformulation*: We explicitly solve the maximization over λ in (5) (Proposition 5.1) to obtain an equivalent smooth convex minimization problem (Theorem 2.8).
   (i) *Strongly convex min-max reformulation*: We propose a new min-max reformulation (21) that is slightly different from (5) and show that it has strong-concavity in λ (Lemma 2.16). This allows us to apply accelerated version of the Chambolle-Pock algorithm (Chambolle and Pock, 2016, Algorithm 4); which converges at a rate of $O(1/k^2)$ in duality gap for ergodic iterates (Remark 2.17)

(b) **Tailored fast algorithm**: We present a novel algorithm (Algorithm 1) called the Fast LIP Solver (FLIPS), which exploits the structure of the proposed smooth reformulation (14) better than the standard acceleration based methods. The novelty of FLIPS is that it combines ideas from canonical gradient descent schemes to find a descent direction; and then from the Frank-Wolfe (FW) algorithm (Jaggi, 2013) in taking a step in the descent direction. We provide an explicit characterization of the optimal step size which could be computed without significant additional computations. We demonstrate the performance of FLIPS on the standard problems of Binary Selection, Compressed sensing, and Image Denoising. particularly for Image Denoising, we show that FLIPS outperforms the state-of-the-art methods for (1) like CP and C-SALSA in both number of iterations and CPUtime.

(c) **Open source Matlab package**: Associated with this algorithm, we also present an open-source Matlab package that includes the proposed algorithm (and also the implementation of CP and C-SALSA) (Sheriff et al., 2022).

This article is organized as follows. In Section 2 we discuss two equivalent reformulations of the LIP; one as a smooth minimization problem in subsection 2.1, and then as a min-max problem with strong-convexity in subsection 2.2. Subsequently, in Section 3 the FLIPS algorithm is presented, followed by the numerical simulations in section 4. All the proofs

of results in this article are relegated towards the end of this article in Section 5 for better readability.

Notations. Standard notations have been employed for the most part. The interior of a set $\mathcal{S}$ as $\mathrm{int}(\mathcal{S})$. The $n \times n$ identity matrix is denoted by $\mathbb{I}_n$. For a matrix M we let $\mathrm{tr}(M)$ and $\mathrm{image}(M)$ denote its trace and image respectively. The gradient of a continuously differentiable function $\eta(\cdot)$ evaluated at a point h is denoted by $\nabla\eta(h)$.

Generally $\|\cdot\|$ is the norm associated with the inner product $\langle \cdot\ ,\ \cdot \rangle$ of the Hilbert space $\mathbb{H}$, unless specified otherwise explicitly. Given two Hilbert spaces $\left(\mathbb{H}_1, \langle \cdot\ ,\ \cdot \rangle_1\right)$ and $\left(\mathbb{H}_2, \langle \cdot\ ,\ \cdot \rangle_2\right)$ and a linear map $T : \mathbb{H}_1 \longrightarrow \mathbb{H}_2$, its adjoint T^a is another linear map $T^a : \mathbb{H}_2 \longrightarrow \mathbb{H}_1$ such that $\langle v\ ,\ T(u) \rangle_2 \ = \ \langle T^a(v)\ ,\ u \rangle_1$ for all $u \in \mathbb{H}_1$ and $v \in \mathbb{H}_2$.

2. Equivalent reformulations with improved convex regularity

We consider the LIP (1) under the setting of following two assumptions that are enforced throughout the article.

Assumption 2.1 (Cost function) *The cost function $c : \mathbb{H} \longrightarrow \mathbb{R}$ is*

(a) positively homogenuous*: For every $r \geqslant 0$ and $f \in \mathbb{H}$, $c(rf) \ = \ rc(f)$*
(b) inf compact*: the unit sub-level set $B_c := \{f \in \mathbb{H} : c(f) \leqslant 1\}$ is compact*
(c) quasi convex*: the unit sub-level set B_c is convex.*

In addition to the conditions in Assumption 2.1, if the cost function is symmetric about the origin, i.e., $c(-f) = c(f)$ for all $f \in \mathbb{H}$, then it is indeed a *norm* on $\mathbb{H}$. Thus, many common choices like the ℓ_1 and Nuclear norms for practically relevant LIPs are included in the setting of Assumption 2.1.

Assumption 2.2 (Strict feasibility) *We shall assume throughout this article that $\|x\| > \epsilon > 0$ and that the corresponding LIP (1) is strictly feasible, i.e., there exists $f \in \mathbb{H}$ such that $\|x - \phi(f)\| < \epsilon$.*

2.1 Reformulation as a *smooth* minimization problem

Let the mapping $e : \mathbb{R}^n \longrightarrow [0, +\infty)$ be defined as

$$e\,(h) := \quad \min_{\theta \in \mathbb{R}} \|x - \theta\phi(h)\|^2 \quad = \quad \begin{cases} \|x\|^2 & \text{if } \phi(h) = 0, \\ \|x\|^2 - \dfrac{|\langle x\ ,\ \phi(h) \rangle|^2}{\|\phi(h)\|^2} & \text{otherwise.} \end{cases} \tag{6}$$

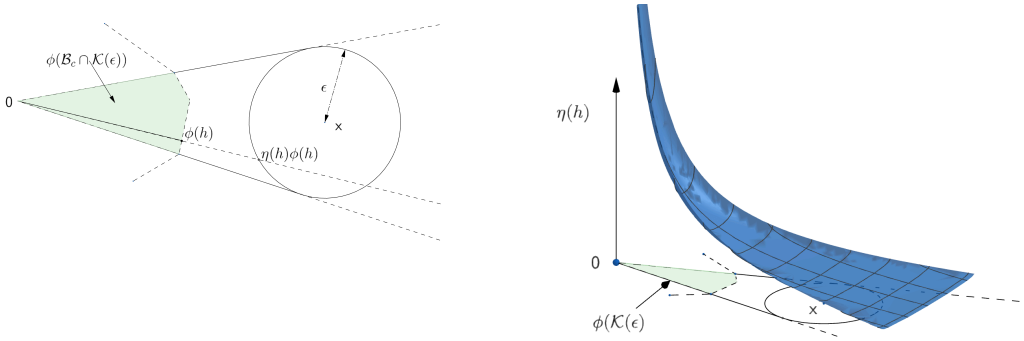
Consider the family of convex cones $\{\mathcal{K}(\bar{\epsilon}) : \bar{\epsilon} \in (0, \epsilon]\}$ defined by

$$\mathcal{K}(\bar{\epsilon}) := \left\{h \in \mathbb{H} : \langle x\ ,\ \phi(h) \rangle \ > \ 0, \text{ and } e\,(h) \leqslant \bar{\epsilon}^2\right\}, \text{ for every } \bar{\epsilon} \in (0, \epsilon]. \tag{7}$$

Equivalently, observe that $h \in \mathcal{K}(\bar{\epsilon})$ if and only if $\langle x\ ,\ \phi\,(h) \rangle \ \geqslant \ \|\phi\,(h)\| \sqrt{\left(\|x\|^2 - \bar{\epsilon}^2\right)}$. Therefore, it is immediately evident that $\mathcal{K}(\bar{\epsilon})$ is convex for every $\bar{\epsilon} \in (0, \epsilon]$.

Definition 2.3 (New objective function) *Consider x, ϕ, and $\epsilon > 0$ such that Assumption 2.2 holds, and let the map $\eta : \mathcal{K}(\epsilon) \longrightarrow [0, +\infty)$ be defined by*

$$\eta(h) := \frac{\|x\|^2 - \epsilon^2}{\langle x\ ,\ \phi(h)\rangle\ +\ \|\phi(h)\|\sqrt{\epsilon^2 - e(h)}}. \tag{8}$$



(a) Diagram presenting the relation between $\phi(B_c \cap \mathcal{K}(\epsilon))$, h, $\eta(h)$, $\phi(h)$, x and ϵ.

(b) Diagram presenting the $\eta(h)$ evaluated over $\phi(\mathcal{K}(\epsilon))$.

Figure 1: Graphical overview of $\eta, \phi\big(\mathcal{K}(\epsilon)\big)$.

Remark 2.4 (Physical interpretation of $\mathcal{K}(\epsilon)$ and $\eta(h)$) For any $h \in \mathbb{H}$, $h \in \mathcal{K}(\epsilon)$ if and only if the ray $\{\theta\phi(h) : \theta \geqslant 0\}$ (i.e., the line going from origin and passing through $\phi(h)$) intersects with $B[x, \epsilon]$. Now, for any $h \in \mathcal{K}(\epsilon)$, the value $\eta(h)$ is the minimum amount by which the point $\phi(h)$ must be scaled so that it intersects with the closed neighborhood $B[x, \epsilon]$ of x as depicted in Figure 1a.

Proposition 2.5 (Derivatives of η) *Consider $x, \phi, \epsilon > 0$ such that Assumption 2.2 holds, and $\eta(h)$ as defined in (8). then the following assertions hold.*

(i) **Convexity**: *The function $\eta : \mathcal{K}(\epsilon) \longrightarrow [0, +\infty)$ is convex.*

(ii) **Gradients**: *The function $\eta : \mathcal{K}(\epsilon) \longrightarrow [0, +\infty)$ is differentiable at every $h \in \operatorname{int}(\mathcal{K}(\epsilon)) = \{h \in \mathbb{H} : \langle x\ ,\ \phi(h)\rangle > 0,\ \text{and } e(h) < \epsilon^2\}$, and the derivative is given by*

$$\nabla\eta(h) = \frac{-\eta(h)}{\|\phi(h)\|\sqrt{\epsilon^2 - e(h)}}\,\phi^a\Big(x - \eta(h)\,\phi(h)\Big) \quad \text{for all } h \in \operatorname{int}(\mathcal{K}(\epsilon)), \tag{9}$$

where ϕ^a is the adjoint operator of ϕ.

(iii) **Hessian**: *the function $\eta : \mathcal{K}(\epsilon) \longrightarrow [0, +\infty)$ is twice continuously differentiable at every $h \in \operatorname{int}(\mathcal{K}(\epsilon))$. The hessian is the linear operator*

$$\mathbb{H} \ni v \longmapsto \big(\Delta^2\eta(h)\big)(v) := (\phi^a \circ M(h) \circ \phi)(v)\ \in\ \mathbb{H}, \tag{10}$$

where $M : \operatorname{int}(\mathcal{K}(\epsilon)) \longrightarrow \mathbb{R}^n \times \mathbb{R}^n$ is a continuous matrix-valued map.[2]

2. Please see (48) for the precise definition of $M(h)$.

Definition 2.6 (Convex Regularity) *Let $\mathcal{H}$ be a convex set and consider $\eta : \mathcal{H} \longrightarrow \mathbb{R}$.*

1. (Smoothness) - *The mapping $\eta : \mathcal{H} \longrightarrow \mathbb{R}$ is said to be β-smooth if there exists $\beta \geqslant 0$ such that the inequality*

$$\left\|\nabla\eta(h) - \nabla\eta(h')\right\| \;\leqslant\; \beta \;\left\|h - h'\right\| \;\text{ holds for all } h, h' \in \mathcal{H}. \tag{11}$$

2. (Strong-Convexity) - *The mapping $\eta : \mathcal{H} \longrightarrow \mathbb{R}$ is said to be α-strongly convex if there exists $\alpha > 0$ such that the inequality*

$$\eta(h') \;\geqslant\; \eta\left(h\right) + \left\langle \nabla\eta(h)\;,\; h' - h\right\rangle + \frac{\alpha}{2}\left\|h' - h\right\|^2 \;\text{ holds for all } h, h' \in \mathcal{H}. \tag{12}$$

Challenges for smoothness: As such, the mapping $\eta : B_c \cap \mathcal{K}(\epsilon) \longrightarrow [0, +\infty)]$ is not smooth in the sense of (11). This is due to two reasons.

1. *High curvature around origin* - Consider any $h \in B_c \cap \mathcal{K}(\epsilon)$, then for $\theta \in (0, 1]$ it is immediate from (8) that $\eta(\theta h) \propto 1/\theta$, and therefore, the mapping $(0, 1] \ni \theta \longmapsto \eta(\theta h)$ is not smooth. Consequently, the mapping $\eta : B_c \cap \mathcal{K}(\epsilon) \longrightarrow \mathbb{R}$ cannot be smooth. As shown in Figure 1b as a simple example, it is easily seen that η achieves arbitrarily large values and arbitrarily high curvature as $\|h\| \longrightarrow 0$.
2. *High curvature at the boundary of $\mathcal{K}(\epsilon)$* - It must be observed that η is not differentiable on the boundary of the cone $\mathcal{K}(\epsilon)$. Moreover, as $e\left(h\right) \uparrow \epsilon^2$, i.e., h approaches the boundary of the cone $\mathcal{K}(\epsilon)$ from its interior, it is apparent from (9) that the gradients of η are unbounded.

It turns out that by avoiding these two scenarios (which will be made more formal shortly), η is indeed smooth over the rest of the set.

Proposition 2.7 (Convex regularity of η) *Consider the LIP in (1) under the setting of Assumption 2.2 with c^* being its optimal value. For every $\widehat{\eta} \geqslant c^*$ and $\bar{\epsilon} \in (0, \epsilon)$, consider the convex set*

$$\mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right) \coloneqq \{h \in B_c \cap \mathcal{K}(\bar{\epsilon}) : \eta\left(h\right) \leqslant \widehat{\eta}\}. \tag{13}$$

(i) **Smoothness:** *There exists constant $\beta > 0$ (see Remark 2.11), such that the mapping $\eta : \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right) \longrightarrow [0, +\infty)$ is β-smooth in the sense of (11).*

(ii) **Strong convexity:** *In addition, if the linear operator ϕ is invertible, then there exists constant $\alpha > 0$ (see Remark 2.12), such that the mapping $\eta : \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right) \longrightarrow [0, +\infty)$ is α-strongly convex in the sense of (12).*

Theorem 2.8 (Smooth reformulation) *Consider the LIP (1) under the setting of Assumption 2.1 and 2.2, and let c^* be its optimal value. Then the following assertions hold*

(i) *Every optimal solution f^* of the LIP (1) satisfies $e(f^*) < \epsilon^2$.*

(ii) **The smooth problem:** *Consider any $(\bar{\epsilon}, \widehat{\eta})$ such that $e(f^*) \leqslant \bar{\epsilon}^2 < \epsilon^2$ and $c^* < \widehat{\eta}$. Then the optimization problem*

$$\min_{h \,\in\, \mathcal{H}(\bar{\epsilon}, \widehat{\eta})} \;\eta\left(h\right). \tag{14}$$

is a smooth convex optimization problem equivalent to the LIP (1). *In other words, the optimal value of* (14) *is equal to* c^* *and* h^* *is a solution to* (14) *if and only if* c^*h^* *is an optimal solution to* (1).

Remark 2.9 (Choosing $\bar{\epsilon}$) Consider the problem of recovering some true signal f^* from its noisy linear measurements $x = \phi(f^*) + \xi$, where ξ is some additive measurement noise. Then, ϵ is chosen in (1) such that the probability $\mathbb{P}\left(\|\xi\| \leqslant \epsilon\right)$ is very high. Then for *any* $\bar{\epsilon} \in (0, \epsilon)$ we have $e(f^*) < \bar{\epsilon}^2$ with probability at least $\mathbb{P}\left(\|\xi\| \leqslant \epsilon\right) \cdot \mathbb{P}\left(\left|\left\langle \xi \ , \ \frac{\phi(f^*)}{\|\phi(f^*)\|} \right\rangle\right|^2 > \epsilon^2 - \bar{\epsilon}^2\right)$. Thus, in practise, one could select $\bar{\epsilon}$ to be just smaller than ϵ based on available noise statistics. For empirical evidence, we have demonstrated in Section 4.6 by plotting the histogram of $e(f^*)$ for 28561 instances of LIPs arising in a single image denoising problem solved with a fixed value of ϵ. It is evident from Figure (7b) that there is a strict separation between the value ϵ^2, and the maximum value of $e(f^*)$ among different LIPs.

Remark 2.10 (Choosing the upper bound $\widehat{\eta}$) Since $\widehat{\eta}$ is any upper bound to the optimal value c^* of the LIP (1), and equivalently (14), a simple candidate is to use $\widehat{\eta} = \eta\left(h\right)$ for any feasible h. In particular, for $h_0 = \left(1/c(f')\right) f'$, where $f' \coloneqq \underset{f}{\operatorname{argmin}} \ \|x - \phi(f)\|^2$ is the solution to the least squares problem, it can be shown that

$$\widehat{\eta} \coloneqq \eta\left(h_0\right) = c(f') \left(1 - \sqrt{1 - \frac{\|x\|^2 - \epsilon^2}{\|x'\|^2}}\right) \quad \text{where } x' = \phi(f'). \tag{15}$$

We would like to emphasise that the value of $\widehat{\eta}$ is required only to conclude smoothness of (14) and the corresponding smoothness constants. It is *not necessary* for the implementation of the proposed algorithm FLIPS. The inequality $\eta\left(h\right) \leqslant \widehat{\eta}$ in the constraint $h \in \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right)$ is ensured for all iterates of FLIPS as it generates a sequence of iterates such that η is monotonically decreasing.

Remark 2.11 (Smoothness parameter) Let $\widehat{\sigma}(\phi^a \circ \phi)$ be the maximum eigenvalue of the linear operator $\left(\phi^a \circ \phi\right)$, then for any $\widehat{\eta} \geqslant c^*$, and $\bar{\epsilon} > 0$ such that $e(f^*) \leqslant \bar{\epsilon}^2 < \epsilon^2$, consider

$$\beta\left(\bar{\epsilon}, \widehat{\eta}\right) \coloneqq \frac{\epsilon^2 \widehat{\eta}^3}{\left(\epsilon^2 - \bar{\epsilon}^2\right)^{3/2}} \frac{\left(\ \|x\| + \epsilon\right)}{\left(\ \|x\| - \epsilon\right)^2} \ \widehat{\sigma}(\phi^a \circ \phi). \tag{16}$$

Then the mapping $\eta : \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right) \longrightarrow [0, +\infty)$ is $\beta\left(\bar{\epsilon}, \widehat{\eta}\right)$-smooth in the sense of (11).

Remark 2.12 (Strong convexity parameter) Let $\bar{\sigma}(\phi^a \circ \phi)$ be the minimum eigenvalue of the linear operator $\left(\phi^a \circ \phi\right)$, then for any $\bar{\eta} \in (0, c^*]$, consider the constant

$$\alpha(\bar{\eta}) = \frac{2\bar{\eta}^3}{\left(\ \|x\|^2 - \epsilon^2\right)} \ \bar{\sigma}\left(\phi^a \circ \phi\right). \tag{17}$$

Then the mapping $\eta : \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right) \longrightarrow [0, +\infty)$ is $\alpha(\bar{\eta})$-strongly convex in the sense of (12). Since c^* is not known a priori, we need a valid lower bound $\bar{\eta}$ that is easy to compute from the problem parameters x, ϵ, and ϕ, which we provide in the following remark.

Remark 2.13 (Choosing the lower bound $\bar{\eta}$) Let $c'(f) := \max_{h \in B_c} \langle f\ ,\ h \rangle$ denote the dual function of c. Then the quantity

$$\bar{\eta} = \frac{\|x\| \left(\|x\| - \epsilon \right)}{c'\big(\phi^a(x)\big)}, \tag{18}$$

is a positive lower bound to the optimal value c^* of the LIP (1).[3]

Remark 2.14 (Applying accelerated gradient descent for (14)) Reformulating the LIP (1) as the smooth minimization problem (14) allows us to apply accelerated gradient descent methods (Ye. E. Nesterov, 1983; Beck and Teboulle, 2009), to improve the theoretical convergence rate from $O(1/k)$ to $O(1/k^2)$. We know that by applying the projected accelerated gradient descent algorithm (Beck and Teboulle, 2009) for (14)

$$\begin{cases} z_k = \Pi_{\mathcal{H}(\bar{\epsilon},\widehat{\eta})}\Big(h_k - (1/b)\nabla\eta(h_k)\Big) \\ t_{k+1} = \dfrac{1+\sqrt{1+4t_k^2}}{2} \\ h_{k+1} = z_k + \dfrac{t_k - 1}{t_{k=1}}(z_k - z_{k-1}), \end{cases} \tag{19}$$

the sub-optimality $\eta(h_k) - c^*$ diminishes at a rate of $O(1/k^2)$, which is an improvement over the existing best rate of $O(1/k)$ for the CP algorithm. Moreover, if the linear map ϕ is invertible, then since the objective function $\eta(c)$ is strongly convex, the iterates in (19) (or even simple projected gradient descent) converge exponentially (with a slightly different step-size rule).

One of the challenges in implementing the algorithm (19) is that it might not be possible in general, to compute the orthogonal projections onto the set $\mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right)$. However, if somehow the inequality $e(h_k) \leqslant \bar{\epsilon}^2$ is ensured always along the iterates, then the problem of projection onto the set $\mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right)$ simply reduces to projecting onto the set B_c, which is relatively much easier. In practice, this can be achieved by selecting $\bar{\epsilon}$ such that $e(f^*) < \bar{\epsilon}^2 < \epsilon^2$, and a very small step-size $(1/b)$ so that the iterates $(h_k)_k$ do not violate the inequality $e(h_k) < \bar{\epsilon}^2$. In our observation, empirically, one can tune the value of b for a given LIP so that the criterion: $e(h_k) < \bar{\epsilon}^2$ is always satisfied. However, if one has to solve a number of LIPs for various values of x via algorithm (19); tuning the value(s) of b could be challenging and tedious. Thus, applying off-the-shelf accelerated methods directly to the smooth reformulation (14) might not be the best choice in practice. This is one of the reasons we propose a different algorithm (FLIPS) that is tailored to solve the smooth reformulation by maximally exploiting the structure of the problem.

3. To see why $\bar{\eta}$ is a valid lower bound, consider (5). We observe that $rx \in \Lambda$ for all $r > 0$, then interchanging the order of min-max in (5), we get

$$c^* \ = \ \max_{\lambda \in \Lambda} \left\{ 2\sqrt{\langle \lambda\ ,\ x \rangle - \epsilon \left\| \lambda \right\|} \ - \ c'(\phi^a(\lambda)) \right\} \ \geqslant \ \max_{r>0} \left\{ 2\sqrt{r}\sqrt{\left\| x \right\|^2 - \epsilon \left\| x \right\|} \ - \ r c'(\phi^a(x)) \right\} \ = \ \bar{\eta}.$$

2.2 Equivalent min-max problem with strong-convexity

Consider the LIP (1) under the setting of Assumptions 2.1 and 2.2, let c^* be its optimal value and f^* be an optimal solution. Consider any $\bar{\eta}, \widehat{\eta}, \bar{\epsilon} > 0$ such that $\bar{\eta} \leqslant c^* < \widehat{\eta}$ and $e(f^*) \leqslant \bar{\epsilon}^2 < \epsilon^2$ (to choose values of $\bar{\epsilon}, \widehat{\eta}, \bar{\eta}$, see Remarks 2.9, 2.10, and 2.13 respectively). Denoting $l(\lambda) \coloneqq \sqrt{\langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\|}$, we define the constant $B > 0$ and the set $\Lambda(\bar{\eta}, B) \subset \mathbb{R}^n$ by

$$\begin{cases} B \coloneqq \dfrac{\epsilon \widehat{\eta}^2}{\big(\|x\|^2 - \epsilon^2 \big) \sqrt{\epsilon^2 - \bar{\epsilon}^2}}, \\ \Lambda(\bar{\eta}, B) \coloneqq \big\{ \lambda \in \mathbb{R}^n : \; l(\lambda) \geqslant \bar{\eta} \text{ and } \; \|\lambda\| \leqslant B \big\}. \end{cases} \tag{20}$$

Lemma 2.15 (Min-max reformulation with strong convexity) *Consider the LIP* (1) *under the setting of Assumptions 2.1 and 2.2, and let c^* denote its optimal value. Then the min-max problem*

$$\begin{cases} \min\limits_{h \in B_c} \; \sup\limits_{\lambda \in \Lambda(\bar{\eta}, B)} \quad L\left(\lambda, h\right) \; = \; 2l(\lambda) \; - \; \langle \lambda \, , \, \phi(h) \rangle, \end{cases} \tag{21}$$

is equivalent to the LIP (1)*. In other words, a pair $(h^*, \lambda^*) \in B_c \times \Lambda(\bar{\eta}, B)$ is a saddle point of* (21) *if and only if $c^* h^*$ is an optimal solution to the LIP* (1)*, and*

$$\lambda^* = \frac{c^*}{\|\phi(h^*)\| \sqrt{\epsilon^2 - e(h^*)}} \big(x - c^* \phi(h^*) \big).$$

The min-max problem (21) falls into the interesting class of convex-concave min-max problems with a bi-linear coupling between the minimizing variable h and the maximizing variable λ. Incorporating the constraints $h \in B_c$ and $\lambda \in \Lambda(\bar{\eta}, B)$ with indicator functions, the min-max problem writes

$$\Big\{ \min_{h \in \mathbb{H}} \; \max_{\lambda \in \mathbb{R}^n} \quad \mathbb{1}_{B_c}(h) \; - \; \langle \lambda \, , \, \phi(h) \rangle \; - \; \big(\mathbb{1}_{\Lambda(\bar{\eta}, B)}(\lambda) - 2l(\lambda) \big). \Big.$$

If the constraint sets B_c and $\Lambda(\bar{\eta}, B)$ are projection friendly, the min-max problem (2.15) can be solved by directly applying the Chambolle-Pock (CP) primal-dual algorithm. Without any further assumptions, the duality gap of min-max problem (21) converges at a rate of $O(1/k)$ for the Chambolle-Pock algorithm. This rate of convergence is currently the best, and same as the one when CP is applied directly to the min-max problem (4) discussed in the introduction. However, in addition, if the mapping $\Lambda(\bar{\eta}, B) \ni \lambda \longmapsto -2l(\lambda)$ is smooth and strongly convex, acceleration techniques can be incorporated into the Chambolle-Pock algorithm. One of the contribution of this article towards this direction is to precisely establish that indeed this mapping is smooth and strongly concave under the setting of Assumption 2.2. in which case, the rate of convergence improves from $O(1/k)$ to $O(1/k^2)$.

Lemma 2.16 (Convex regularity in min-max reformulation) *Consider $x \in \mathbb{R}^n$ and $\epsilon > 0$ such that $\|x\| > \epsilon$. For any $\bar{\eta}, \widehat{\eta}, \bar{\epsilon} > 0$ such that $\bar{\eta} \leqslant c^* < \widehat{\eta}$ and $e(f^*) \leqslant \bar{\epsilon}^2 < \epsilon^2$; let $B > 0$ be as given in* (20)*, and let α', β' be constants given by*

$$\alpha' \coloneqq \frac{\epsilon}{\big(\|x\|^2 - \epsilon^2 \big)} \left(\frac{\bar{\eta}}{B} \right)^3 \qquad \text{and} \qquad \beta' \coloneqq \frac{\big(\|x\|^2 - \epsilon^2 \big)}{2 \bar{\eta}^3}. \tag{22}$$

Then the mapping $\Lambda(\bar{\eta}, B) \ni \lambda \longmapsto -2l(\lambda)$ *is* α'*-strongly convex and* β'*-smooth.*

The Accelerated Chambolle-Pock algorithm. In view of the Lemmas 2.15 and 2.16, the min-max problem (21) admits an accelerated version of the Chambolle-Pock algorithm (Chambolle and Pock, 2016, Algorithm 4, (30)). Denoting Π_{B_c} and $\Pi_{\Lambda(\bar{\eta},B)}$ to be the projection operators onto the sets B_c and $\Lambda(\bar{\eta}, B)$ respectively, and $\big(\alpha', \beta'\big) = \big(\alpha'(\bar{\eta}, B), \beta'(\bar{\eta})\big)$ for simplicity, the accelerated CP algorithm for (21) is

$$\begin{cases} h_{k+1} = \Pi_{B_c}\Big(h_k \; + \; s_k\phi^a\big(\lambda_k + \theta_k(\lambda_k - \lambda_{n-1})\big)\Big) \\ \lambda_{k+1} = \Pi_{\Lambda(\bar{\eta},B)}\left(\lambda_k \; + \; \dfrac{t_k}{l(\lambda_k)}\Big(x - (\epsilon/\|\lambda_k\|)\lambda_k + l(\lambda_k)\phi(h_{k+1})\Big)\right), \end{cases} \tag{23}$$

where, (t_k, s_k, θ_k) are positive real numbers satisfying

$$\theta_{k+1} = \frac{1}{\sqrt{1+\alpha' t_k}}, \quad t_{k+1} = \frac{t_k}{\sqrt{1+\alpha' t_k}}, \quad \text{and} \quad s_{k+1} = s_k\sqrt{1+\alpha' t_k}, \quad \text{for } n \geqslant 0, \tag{24}$$

with $\theta_0 = 1,\ t_0 = \frac{1}{2\beta'}$, and $s_0 = \frac{\beta'}{\|\phi\|_o^2}$.

Remark 2.17 (Ergodic $O(1/k^2)$ rate of convergence) Consider the min-max problem (21), and let (h_k, λ_k) for $n = 1, 2, \ldots,$ be the sequence generated by the Accelerated CP algorithm (23). Then there exists a constant $C > 0$ such that

$$\max_{\lambda\in\Lambda(\bar{\eta},B)} L(h_k, \lambda) \; - \; \min_{h\in B_c} L(h, \lambda_k) \; \leqslant \; \frac{C}{k^2} \quad \text{for all } n \geqslant 1.$$

The remark is an immediate consequence of Lemma 2.16 and (Chambolle and Pock, 2016, Theorem 4 and Lemma 2).

Remark 2.18 (Projection onto the set $\Lambda(\bar{\eta}, B)$) In general, computing projections onto the set $\Lambda(\bar{\eta}, B)$ is non-trivial, and in principle, requires a sub problem to be solved at each iteration. However, since the duality gap along the iterates (h_k, λ_k) generated by (23) converges to zero; it follows that $c^* = \lim\limits_{k\longrightarrow+\infty} l(\lambda_k)$. By selecting $\bar{\eta} < c^*$, computing projections onto the set $\Lambda(\bar{\eta}, B)$ becomes trivial for all but finitely many iterates in the sequence $(\lambda_k)_k$. To see this, observe that the set $\Lambda(\bar{\eta}, B)$ is the intersection of two convex sets $\{\lambda : \|\lambda\| \leqslant B\}$ and $\{\lambda : l(\lambda) \geqslant \bar{\eta}\}$. Since $c^* = \lim\limits_{k\longrightarrow+\infty} l(\lambda_k)$, the inequality $l(\lambda_k) \geqslant \bar{\eta}$ is readily satisfied for all but finitely many iterates if $\bar{\eta} < c^*$. Consequently, all but finitely many iterates in the sequence $(\lambda_k)_k$ are contained in the set $\{\lambda : l(\lambda) \geqslant \bar{\eta}\}$. Therefore, computing projections onto the set $\Lambda(\bar{\eta}, B)$ eventually reduces to projecting onto the set $\{\lambda : \|\lambda\| \leqslant B\}$, which is trivial.

3. The Fast LIP Solver (FLIPS)

Even though the newly proposed smooth reformulation of (1) is amenable to acceleration based schemes; it could suffer in practice from conservative estimation of smoothness constant. Moreover, since one has to ensure that $h \in \mathcal{K}(\epsilon)$ for all the iterates, it further

constrains the maximum step-size that could be taken, which results in slower convergence in practice. These issues make applying off-the shelf methods to solve the proposed smooth problem not fully appealing. To overcome this, we propose the Fast LIP Solver (FLIPS), presented in Algorithm 1.

Algorithm 1: The Fast LIP Solver

Input: Measurement x, linear operator ϕ, $\epsilon > 0$, and oracle parameters.
Output: Sparse representation f^*
Initialise: $h = \big(1/c(f')\big) f'$, where $f' = \phi \backslash x := \operatorname{argmin}_f \|x - \phi(f)\|$.
Check for strict feasibility (Assumption 2.2)
Iterate till convergence

1 **Compute $\eta(h)$ and $\nabla\eta(h)$**

2 **Check for stopping criteria (for small enough $\delta \sim 0.01$)**

$$\text{Stopping criterion}: \quad \frac{\langle \nabla\eta(h)\ ,\ h\rangle}{\min\limits_{g\in B_c} \langle \nabla\eta(h)\ ,\ g\rangle} \geqslant\ 1+\delta. \tag{25}$$

3 **Compute the update direction $g(h)$ using any viable oracle**

4 **Exact line search:** Compute

$$\gamma(h) = \begin{cases} \operatorname{argmin}\limits_{\gamma\in[0,1]} & \eta\big(h+\gamma(g(h)-h)\big) \\ \text{subject to} & h+\gamma(g(h)-h)\ \in\ \mathcal{K}(\bar{\epsilon}) \end{cases}$$

5 **Update** : $h^+ = h + \gamma(h)\big(g(h)-h\big)$
Repeat

6 **Output:** the sparse representation $f^* = \eta(h)h$.

In a nutshell, FLIPS uses two oracle calls in each iteration; one each to compute an update direction $g(h)$ and the step-size $\gamma(h)$. It then updates the current iterate by taking the convex combination $h^+ = h + \gamma(h)\big(g(h)-h\big)$ controlled by $\gamma(h)$. This is repeated until a convergence criterion is met.

Remark 3.1 (Initialization and checking feasibility) The algorithm is initialised with a normalized solution of the least squares problem: $\operatorname{argmin}_f \|x-\phi(f)\|$, which is written as $\phi\backslash x$ following the convention used in Matlab. If the LIP is ill-posed, the least squares problem will have infinitely many solutions. Even though the algorithm works with any initialization among the solutions to the least squares problem, it is recommended to use the minimum ℓ_2-norm solution $f' = \phi^\dagger x$. Since $\phi(f')$ is closest point (w.r.t. the $\|\cdot\|$ used in the least squares problem), it is easily verified that the LIP satisfies the strict feasibility condition in Assumption 2.2 if and only if the inequality $\|x-\phi(f')\| < \epsilon$ holds.

Remark 3.2 (Constraint splitting and successive feasibility) The novelty of FLIPS is that it combines ideas from canonical gradient descent methods to compute the update

direction, but takes a step in spirit similar to that of the Frank-Wolfe algorithm. This allows us to perform a sort of constraint splitting and handle different constraints in (14) separately. To elaborate, recall that the feasible set in (14) is

$$\mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right) = \{h \in B_c \cap \mathcal{K}(\bar{\epsilon}) : \eta\left(h\right) \leqslant \widehat{\eta}\}.$$

On the one hand, since B_c is a convex set, for a given $h \in B_c$, the direction oracle guarantees that $h^+ \in B_c$ by producing $g(h) \in B_c$ at every iteration. On the other hand, selection of $\gamma(h)$ in the exact line search oracle ensures that $\eta(h^+) \leqslant \eta\left(h\right) \leqslant \widehat{\eta}$ and $h^+ \in \mathcal{K}(\bar{\epsilon})$. Thus, $h^+ \in \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right)$, and

3.1 Descent direction oracle

For any given $h \in \mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right)$ (in principle, for any $h \in \operatorname{int}\mathcal{K}(\epsilon)$), the direction oracle simply computes another point $g(h) \in B_c$ such that the function $\eta(\cdot)$ could be potentially minimized along the direction $g(h) - h$. To find $g(h)$, a sub-problem is solved at every iteration of the algorithm. Thus, by varying the complexity of these sub-problems, gives rise to different direction oracles. In addition, the output of a descent direction oracle also provides access to quantities that can be used to define the stopping criteria for FLIPS. In the following, we briefly describe some standard descent direction oracles that could be used in FLIPS along with the corresponding stopping criteria for them.

(a) **Linear Oracle (LO):** For any $h \in \operatorname{int}\mathcal{K}(\epsilon)$, the Linear oracle computes the direction $g(h)$ by solving a linear optimization problem over B_c

$$\text{LO:} \quad g(h) \in \left\{ \underset{g \in B_c}{\operatorname{argmin}} \quad \langle \nabla\eta(h)\ ,\ g \rangle \right. . \tag{26}$$

Finding $g(h)$ via a linear oracle makes the corresponding implementation of FLIPS very similar to the Frank-Wolfe (FW) algorithm (Frank and Wolfe, 1956; Jaggi, 2013), but with constraint splitting as discussed in remark 3.2. The only difference between the FW-algorithm and FLIPS is that the linear sub-problem (26) is solved over the set B_c in FLIPS and not over the actual feasible set $\mathcal{H}\left(\bar{\epsilon}, \widehat{\eta}\right)$ as we would in the FW-algorithm.

Example 1 (LO for sparse coding) For the sparse coding problem, i.e., LIP (1) with $c(f) = \|f\|_1$, the corresponding linear oracle (26) is easily described due to the Hölder inequality (Yang, 1991). The i-th component $g_i(h)$ of the direction $g(h)$ is given by

$$g_i(h) = \begin{cases} -\operatorname{sgn}\left(\partial\eta/\partial h_i\right), & \text{if } |\partial\eta/\partial h_i| = \|\nabla\eta(h)\|_\infty, \\ 0, & \text{if } |\partial\eta/\partial h_i| \neq \|\nabla\eta(h)\|_\infty. \end{cases} \tag{27}$$

(b) **Simple Quadratic Oracle (SQO)**: For any $h \in \operatorname{int}\mathcal{K}(\epsilon)$, the SQO computes the direction $g(h)$ by solving a quadratic optimization problem over B_c.

$$\text{SQO:} \quad g(h) = \Pi_{B_c}\Big(h - (1/\beta)\nabla\eta(h)\Big) \tag{28}$$

where Π_{B_c} is the projection operator onto the set B_c. Thus, an SQO is essentially a composition of taking a gradient descent step with a step-size of $1/\beta$ and projecting

back to the set B_c. The parameter β is a hyper parameter of the SQO, which is usually taken as the inverse of smoothness constant in canonical gradient descent schemes. We emphasise here that FLIPS does not jump from h to $g(h)$ right away as in projected gradient descent algorithm, but rather takes a convex combination of these points. This allows us to chose β much smaller than the actual smoothness constant.

(c) **Accelerated Quadratic Oracle (AQO)**: Adding momentum/acceleration in gradient descent schemes tremendously improves their convergence speeds, both in theory and practice (Ye. E. Nesterov, 1983). Taking inspiration from such ideas, we consider the AQO as

$$\text{AQO:} \quad \begin{cases} g(h,d) = \Pi_{B_c}\Big(h - (1/\beta)\big(\nabla\eta(h) + \rho d\big)\Big) \\ d(h,d) = g(h,d) - h \end{cases} \tag{29}$$

The extra iterate d carries the information of the momentum/past update, and the parameter ρ controls the weight of the momentum, which is an additional hyper parameter of the AQO.

Solving the quadratic problems (28) and (29) require more computational resources than the linear one (26). Consequently, the complexity of implementing a quadratic oracle is more than that of the linear oracle. Since, solving the quadratic problem (28) reduces to computing orthogonal projections of points onto the set B_c, a practical assumption in implementing a quadratic oracle is that the set B_c is projection friendly, which is indeed the case whenever the corresponding cost function $c(\cdot)$ is *Prox-friendly*. For some LIPs like the matrix completion problem, solving the corresponding projection problem requires computing the SVD at every iteration, which could be challenging for large scale problems. However, for other relevant cost functions like the ℓ_1, ℓ_∞-norms, the corresponding projection problem requires projection onto the corresponding spheres which is easy to implement, for e.g., (Duchi et al., 2008).

3.2 Step size selection via exact line search

Once the direction $g(h)$ is computed using a viable oracle, FLIPS updates the iterate h by taking a convex combination: $h + \gamma(g(h) - h)$ in spirit similar to that of Frank-Wolfe algorithm. We select the step-size $\gamma \in [0,1]$ via exact line search, i.e., by solving the problem

$$\gamma(h) = \operatorname*{argmin}_{\gamma\in[0,1]} \quad \eta\big(h + \gamma(g(h) - h)\big). \tag{30}$$

Luckily, the reformulation of the LIP as (14) with new objective function η allows us to compute the explicit solution of (30) without any noticeable increase in the computational demand. The following proposition characterises the optimal step-size in the exact line search for a generic direction d instead on specific $g(h) - h$ as in (30).

Proposition 3.3 *For any $h \in \mathcal{K}(\epsilon)$, and $d \in \mathbb{H}$, then consider the optimization problem*

$$\gamma^* = \operatorname*{argmin}_{\gamma\in[0,1]} \quad \eta(h + \gamma d). \tag{31}$$

Then exactly one of the following assertions hold

(1) *If* $\langle x\ ,\ \phi(d)\rangle \leqslant \eta\,(h)\,\langle \phi(h)\ ,\ \phi(d)\rangle$, *then* $\gamma^* = 0$.
(2) *If* $h + d \in \mathcal{K}(\epsilon)$ *and* $\langle x\ ,\ \phi(d)\rangle \geqslant \eta(h+d)\,\langle \phi(h+d)\ ,\ \phi(d)\rangle$, *then* $\gamma^* = 1$.
(3) *Otherwise,* $\gamma^* \in (0,1)$ *is the root of the quadratic equation* $a\gamma^2 + 2b\gamma + c = 0$,[4] *that also satisfies*

$$\frac{\langle x\ ,\ \phi(h+\gamma^* d)\rangle}{\left(\|x\|^2 - \epsilon^2\right)} \leqslant \frac{\langle \phi(h+\gamma^* d)\ ,\ \phi d\rangle}{\langle x\ ,\ \phi(d)\rangle}.$$

Remark 3.4 (Qualitative convergence of FLIPS) *We would like to provide qualitative convergence of FLIPS with a 'Simple Quadratic Oracle (SQO)'. Consider the sequence of iterates* $(h_k)_k$*, generated form FLIPS with an SQO, then the following arguments apply*

(i) *the sequence of the values of the cost function* $(\eta(h_k))_k$ *is non-increasing for FLIPS with any oracle, and in particular, with an SQO*
(ii) *the mapping* $h_t \longmapsto \big(g(h_k), \gamma(h_k)\big)$*, is continuous for all* t
(iii) *the equality* $h = g(h)$ *holds if and only if* $h = h^*$

Claim (iii) is non-trivial but follows directly from the first order optimality conditions, we omit this proof for the sake of brevity. Putting the three arguments (i) - (iii) together, and considering the Lyapunov function $V(h) := \eta(h) - \eta(h^*)$*, we conclude from the Lyapunov theorem that the iterations* $(h_k)_k$ *converge to* h^*.

Remark 3.5 (Techniques for speeding up implementation of FLIPS) We discuss some techniques and tricks to improve the practical implementation of FLIPS

1. The inner-product terms $\langle x\ ,\ \phi\,(h)\rangle, \langle x\ ,\ \phi(d)\rangle$ must be computed as $\langle \phi^a(x)\ ,\ h\rangle$ and $\langle \phi^a(x)\ ,\ d\rangle$. This way, we do not compute (and store) the vectors $\phi(h)$ and $\phi(d)$ at each iteration, but rather compute the vector $\phi^a(x)$ once.
2. To compute $\nabla\eta(h)$, we need to compute $\phi^a(\lambda(h))$. This can be alternatively done by keeping a running iterate of $(\phi^a o \phi)h$, which is updated at each iteration as $(\phi^a o\phi)h_{t+1} = (\phi^a o\phi)h_t + \gamma_t\big(\phi^a o\phi(g_t) - (\phi^a o\phi)h_t\big)$.

4. Numerical Results

We test FLIPS on a few well-known LIPs namely *Compressed Sensing, Binary-Selection problem*, and finally we test FLIPS on the classical *Image Denoising* problem. Moreover, since the image processing problems are the most common LIPs and many good solvers already exist, we compare FLIPS with existing state of the art methods for constrained LIPs arising in image processing tasks.

4.

$$\begin{aligned} a &= \|\phi(d)\|^2\,(e\,(d) - \epsilon^2) \\ b &= (\|x\|^2 - \epsilon^2)\langle \phi(h\ ),\ \phi(d)\rangle - \langle x\ ,\ \phi(h)\rangle\langle x\ ,\ \phi(d)\rangle \\ c &= \frac{(\|x\|^2 - \epsilon^2)\langle \phi(h)\ ,\ \phi(d)\rangle^2}{\|\phi(d)\|^2} - \frac{2\langle x\ ,\ \phi(d)\rangle\langle x\ ,\ \phi(h)\rangle\langle \phi(h)\ ,\ \phi(d)\rangle}{\|\phi(d)\|^2} + \frac{\|\phi(h)\|^2\langle x\ ,\ \phi(d)\rangle^2}{\|\phi(d)\|^2}. \end{aligned} \tag{32}$$

All experiments in Sections 4.2, 4.1, and 4.3 were run on a laptop with Apple M1 processor with 8GB RAM using MATLAB 2021b. While, the experiments in Section 4.4 comparing FLIPS with FISTA were run on a laptop with Apple M1 max processor with 32GB RAM using Python. The open-source code can be found on the author's Github page (Sheriff et al., 2022).

4.1 Results on the Binary-selection problem

In this example, we aim to reconstruct a vector $f_{tr} \in \mathbb{R}^K$ whose entries are ± 1 from its linear measurements. Without loss of generality, we consider $f_{tr}(i) = +1$ for $i = 1, 2, \ldots, 0.5K$, and $f_{tr}(i) = -1$ for $i > 0.5K$, and then collect $m \sim 0.55K$ linear measurements $x \in \mathbb{R}^m$ (55% of the information). For each $i = 1, 2, \ldots, m$, the measurement x_i is obtained as $x_i = \phi_i^\top f_{tr} + w_i$, where $\phi_i \in \mathbb{R}^K$ is generated randomly by sampling each entry of ϕ_i uniformly over the interval $[-0.5, 0.5]$. The measurement noise $w_i \sim \mathcal{N}(0, \sigma)$ with $\sigma = 0.0125$, is generated by sampling randomly and independently from everything else.

Following, Chandrasekaran et al. (2012), the problem of recovering f_{tr} from x, ϕ is formulated as the LIP

$$\begin{cases} \min\limits_{f} & \|f\|_\infty \\ \text{subject to} & \|x - \phi f\| \leqslant \epsilon. \end{cases} \tag{33}$$

We chose the value of $\epsilon = 10\sigma\sqrt{m}$. Since we know that the entries can be either $+1$ or -1, it allows us to select a slightly larger value of ϵ than necessary, which has an indirect benefit in improving the regularity of the problem and consequently faster convergence.

We consider three different instances of (33) for $K = 500, 1000$, and 5000. We apply FLIPS for each of the problems using an AQO oracle with parameters β and ρ tuned for faster convergence. The performance of FLIPS for $K = 3000$ iterations is shown in Figure 2 following the theme of Figure 3. The first row plots the true signal f_{tr} and the recovered signal $f^* = \eta(h_T)h_T$. Following these to the bottom we have the plots for the *sub-optimality*, *Distance to true solution*, and the sequence of *step-sizes* as a function of iterations of FLIPS.

4.2 Results on Compressed Sensing

For an image 'I', let $\mathsf{I} \in \mathbb{R}^K$ be its vectorized form. Then for $i = 1, 2, \ldots, m \sim 0.6 \times K$, we collect the linear measurement x_i, of the image 'I' as $x_i = c_i^\top \mathsf{I} + w_i$; where $c_i \in \mathbb{R}^K$ is a random vector whose each entry is drawn uniformly from $[-0.5, 0.5]$, and w_i is the measurement noise drawn randomly from a Gaussian distribution with variance $\sigma^2 = 0.0055$, and independently from everything else. The task of recovering the image 'I' from its measurements x is formulated as the LIP

$$\begin{cases} \min\limits_{f} & \|f\|_1 \\ \text{subject to} & \|x - (CD)f\| \leqslant \epsilon, \end{cases} \tag{34}$$

where $\epsilon = \sigma\sqrt{m}$ and $D \in \mathbb{R}^{K \times K}$ is chosen to be the dictionary of 2d-IDCT basis vectors since natural images are sparse in 2d-DCT basis. Thus, (34) is a version of the LIP (1) with

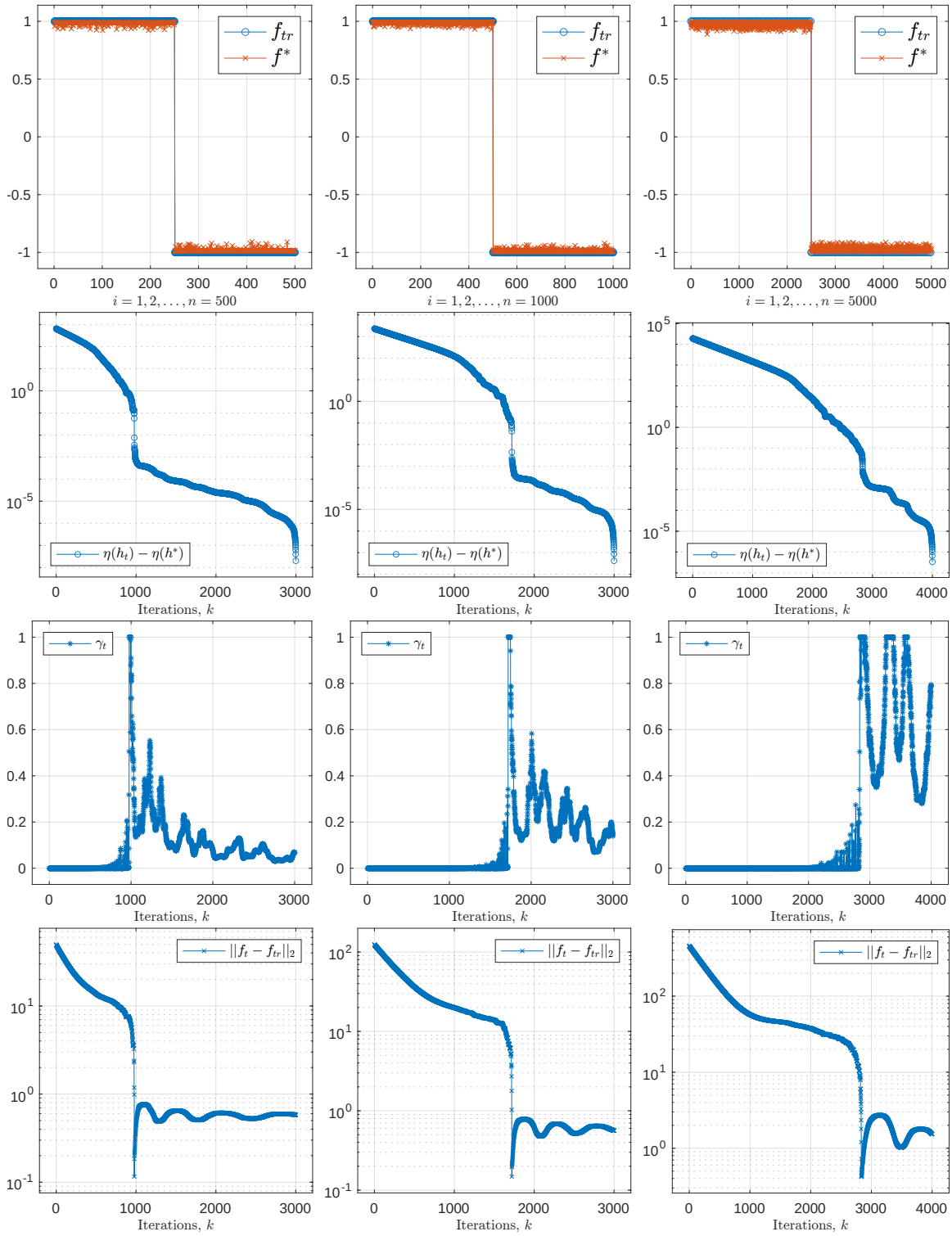


Figure 2: Simulation results for Binary selection.

objective function $c(\cdot) = \|\cdot\|_1$ and parameters $x, \phi = CD, \epsilon$. We apply FLIPS with an AQO to find an optimal solution f^* to the LIP (34), and the image 'I' is reconstructed as Df^*.

We consider the Compressed Sensing problem (34) for the standard images of 'Lena', 'Cameraman', and 'Barbara', each of size $K = 64 \times 64$ pixels. We solve each one of them using FLIPS using an AQO oracle with parameter values tuned for faster convergence. The results on recovery and convergence attributes of FLIPS are shown in Figure 34. The left column corresponds to the results for the 'Lena' image, followed by 'Barbara', and 'Cameraman' images to their right. The true images are shown in the first row and the recovered images from FLIPS are shown in the second row. Following these, the plots for the *sub-optimality*: $\eta(h_t) - \eta(h^*)$, *Distance to true solution*: $\|f_t - \mathrm{I}\|_2$ is shown where $f_t = \eta(h_t)\, h_t$. Finally, at the bottom, we plot the sequence of *step-sizes*: γ_t as a function of iterations of FLIPS.

4.3 Results on image denoising

Finally, we also consider another image processing task of denoising an image to demonstrate the performance of FLIPS and compare it with other state-of-the-art methods that denoise an image by solving the corresponding constrained LIP (1). In particular, we consider the Chambolle-Pock algorithm (with the current best theoretical convergence guarantee of $O(1/k)$) (Chambolle and Pock, 2016, 2010), and also the more well known C-SALSA algorithm (Afonso et al., 2009).

Table 1: Comparison of FLIPS with CP and C-SALSA algorithms, for image denoising with sliding patches on the 'Lena', 'Barbara', and 'Cameraman' images.

	FLIPS		**CP**		**C-SALSA**	
	CPU time	iterations	CPU time	iterations	CPU time	iterations
Results for "Lena"						
4 x 4	4.95	5.188	14.54	45.18	16.9	43.61
8 x 8	5.6	4.9	17.7	45.73	14.27	34.33
16 x 16	16	3	33.85	47.6	27.8	44.5
32 x 32	27.23	2.6	70.23	49	43.2	43.2
Results for "Barbara"						
4 x 4	5.46	5.37	15.1	46.2	15.95	44.1
8 x 8	5.74	5.08	18.13	47.66	14.5	36.3
16 x 16	16	3	33.67	49.1	27.58	47.5
32 x 32	27.6	2.6	71.5	49.7	44.5	43
Results for "Cameraman"						
4 x 4	5.3	5.33	13.73	44.4	16.28	43.9
8 x 8	5.92	5.313	17.93	43.5	12.63	31.18
16 x 16	15.9	3	30.45	41.75	25.09	38.7
32 x 32	28.4	2.88	66.78	46.88	42.96	42.69

We consider the three images: 'Lena', 'Barbara', and 'Cameraman', each of size 256×256. On each of those images, a Gaussian noise of variance $\sigma^2 = 0.0055$ is added with the

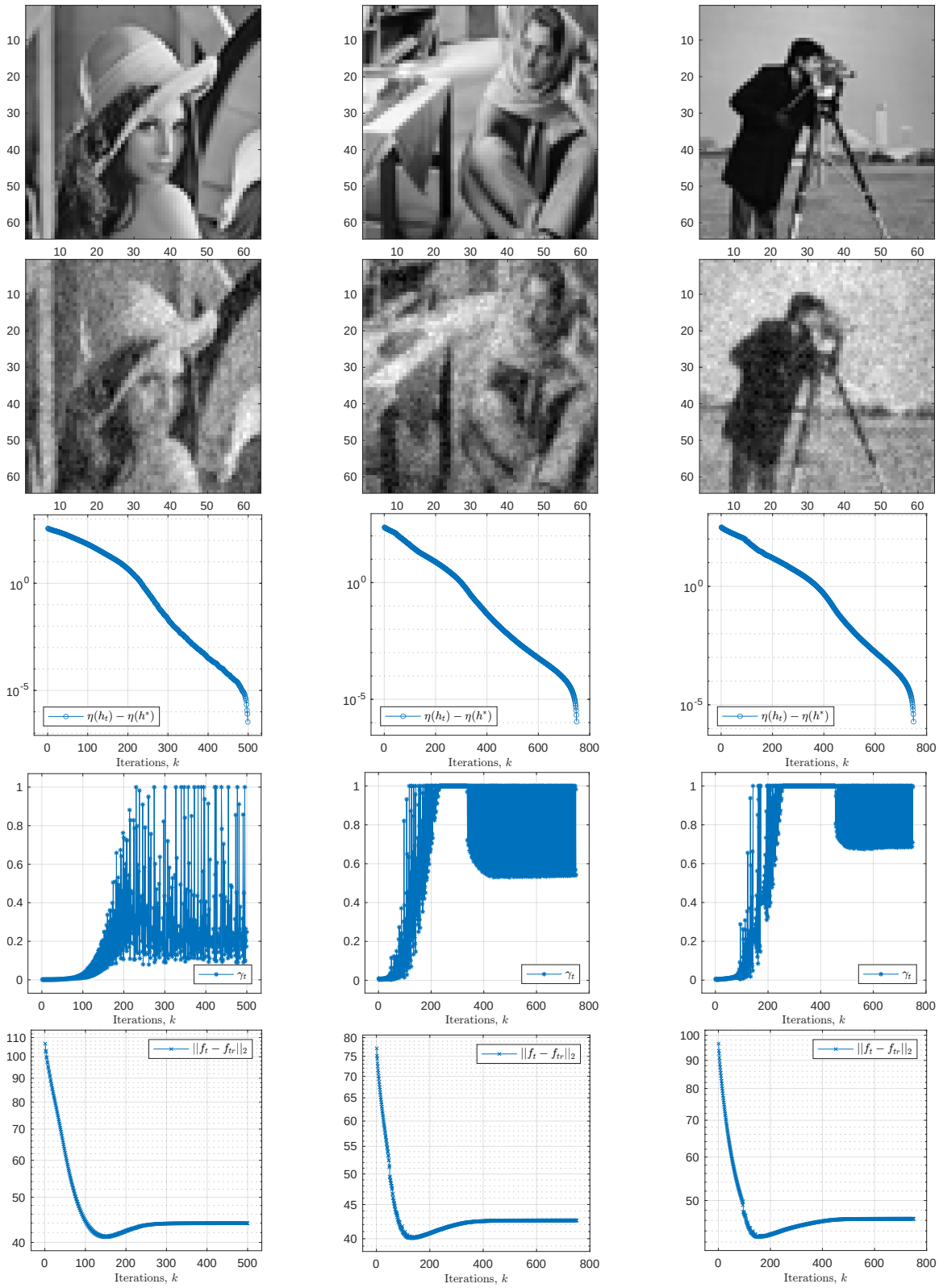


Figure 3: Simulation results for Compressed sensing.

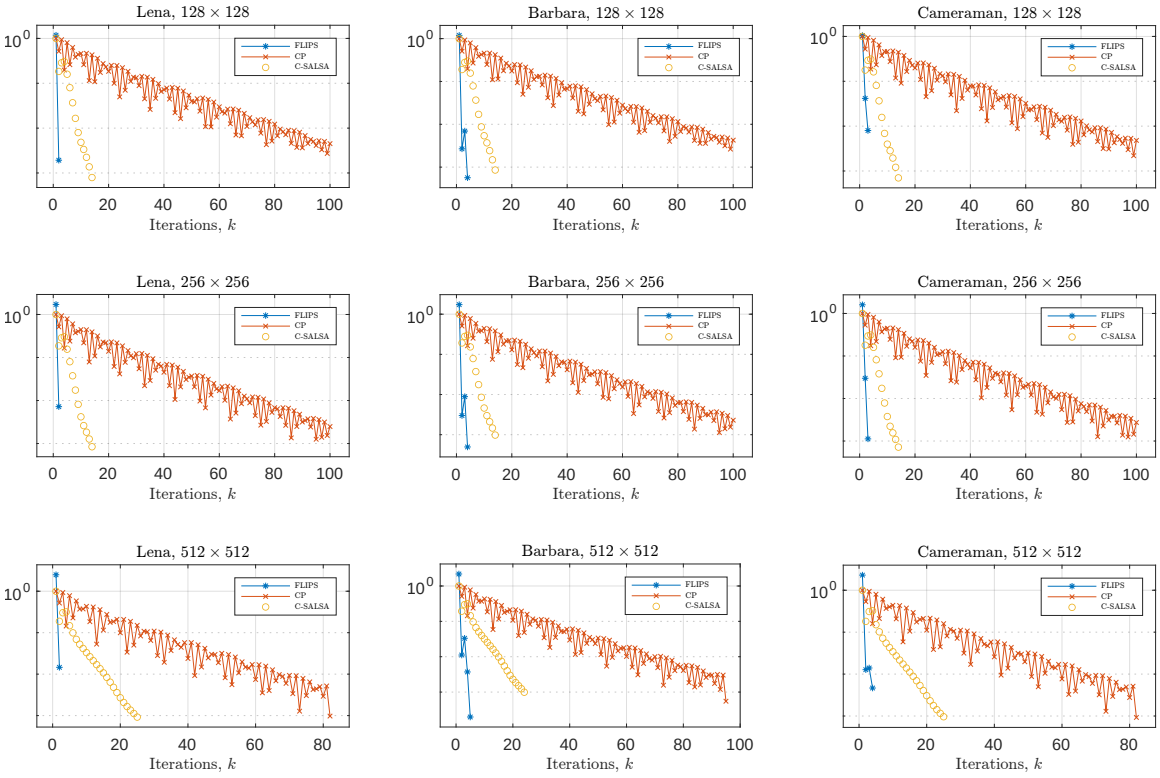


Figure 4: Comparison of FLIPS with C-SALSA and CP algorithms on image denoising (full images).

MATLAB function `imnoise`. The image is then denoised by denoising every patch (and overlapping) of a fixed size $m \times m$. Then the final image is reconstructed by taking the average value of an individual pixel over all the patches it belongs to. Denoising a given patch of size $m \times m$ corresponds to solving the LIP (1) with $c(f) = \|f\|_1$, the linear map ϕ as the 2d-inverse discrete cosine transform for $m \times m$ patches (computed using the function `idct2` in Matlab), and $\epsilon = \sigma m$. The experiment is repeated with different patch sizes of $m = 4, 8, 16, 32, 64$, and for each image and patch size m, the convergence results are averaged over 10 independent trials with independent noise. The parameter values in the Chambolle-Pock algorithm, C-SALSA, and the AQO oracle parameters in FLIPS are tuned to get the best results for each patch size.

In Table 1, we tabulate the average number of iterations per patch required until convergence for each algorithm, averaged over different patches of the respective image and different instances of noise. Besides the average number of iterations, we also tabulate the total CPU time each algorithm takes to solve the denoising problem for all the patches, and averaged over different instances of noise. The convergence criterion for each patch I is chosen to be $\min\left\{k : \|\mathsf{I}_k - \mathsf{I}^*\|_2 \leqslant 10^{-3}\,\|\mathsf{I}^*\|_2\right\}$, where I^* is the optimal solution for the respective patch computed apriori by running FLIPS for a large number of iterations. Of course, the time required to compute $\|\mathsf{I}_k - \mathsf{I}^*\|_2$ at each iteration is excluded from the CPU times.

Table 2: Comparison of FLIPS with CP and C-SALSA algorithms, for image denoising on full images of 'Lena', 'Barbara', and 'Cameraman' images.

	FLIPS		CP		C-SALSA	
	CPU time	# iteration	CPU time	# iteration	CPU time	# iteration
Results for "Lena"						
128 x 128	0.023	2	0.163	60	0.027	9
256 x 256	0.045	2.25	0.275	53	0.07	9
512 x 512	0.219	3	0.54	42	0.39	16
Results for "Barbara"						
128 x 128	0.035	2	0.145	53	0.036	9
256 x 256	0.043	2	0.285	53	0.068	9
512 x 512	0.286	4	0.627	53	0.347	15
Results for "Cameraman"						
128 x 128	0.026	3	0.137	55	0.025	9
256 x 256	0.06	3	0.289	53	0.073	9
512 x 512	0.248	4	0.525	42	0.388	16

In addition, we also consider the denoising problem for the three images by directly working on the entire image as a single patch instead of considering smaller and sliding patches as in Table 1. For this, we first obtain noise-free images of size 128×128, 256×256, and 512×512 pixels. Then, similar to the previous experiment, a Gaussian noise of variance $\sigma^2 = 0.0055$ is added with the MATLAB function `imnoise` to obtain the noisy image. Then the noisy image is denoised by solving the corresponding LIP (for the full image) using FLIPS, CP, and C-SALSA algorithms with respective parameters tuned to give better respective convergence results.

We compare their convergence attributes on the metric $\|\mathsf{I}_k - \mathsf{I}^*\|$, where I_k is the image after k iterations of the respective algorithm and I^* is the optimal solution obtained apriori by running FLIPS for many iterations (and confirmed with other methods for optimality). Convergence plots for FLIPS, CP, and C-SALSA algorithms for image denoising on the three images of 'Lena', 'Barbara', and 'Cameraman' for varying sizes of 128×128, 256×256, and 512×512 are provided in Figure 4, and the corresponding CPU times (averaged over 10 iterations of different noise) is tabulated in Table 2. It must be observed that FLIPS only takes approximately ~ 4 iterations to converge to the optimal solution, which is incredibly fast.

4.4 Comparison with FISTA for a trajectory of solutions

Finally, we would like to compare the convergence of FLIPS with FISTA for solving an LIP. In this regard, we first generate the problem data, namely: (i)-liner map ϕ, (ii)-Sparse vectors F^{tr}, and (iii)-noisy measurements X, randomly and independently from each other as

1. $\phi \in \mathbb{R}^{m \times K}$ is generated randomly by sampling each entry $\phi_{ij} \sim \mathcal{N}(0, 1)$

2. $F^{tr} \in \mathbb{R}^{K\times T}$, is randomly generated by first sampling every entry $F^{tr}_{ij} \sim \mathcal{N}(0,1)$, and then every column is made to be S-sparse by zeroing all but the S-largest entries in magnitude
3. We first obtain the true measurements as $\mathbb{R}^{m,N} \ni X^{tr} = \phi \cdot F^{tr}$, from which the noisy measurements $X = X^{tr} + \sigma_w W$ are obtained by adding an AWGN $W \in \mathbb{R}^{m\times T}$. Each entry $W_{ij} \sim \mathcal{N}(0,1)$ is iid and sampled independently from the previous data (i.e., ϕ and F^{tr}), and $\sigma_w > 0$ is a scalar constant chosen to satisfy a specified SNR level as

$$\sigma_w = \left\|X^{tr}\right\|_{fro} \sqrt{\left(1/mN\right) 10^{\frac{-SNR}{10}}}$$

Given the problem data: (ϕ, X), we obtain two estimates of F^{tr} by solving two different formualtions of a Linear Inverse problem. Let $F^c(\epsilon)$ be the estimate obtained by solving the *constrained* formulation of the LIP using FLIPS, and $F^r(\lambda)$ be the estimate obtained by solving the ℓ_1-*regularized LASSO* formulation using FISTA. To this end, for any $\epsilon > 0$ and $\lambda > 0$, let us define

$$\begin{cases} F^c(\epsilon) \in \underset{F}{\operatorname{argmin}} & \|F\|_{(1,1)} \quad \text{subject to } \|X - \phi F\|_{fro} \leqslant \epsilon, \text{ and} \\ F^r(\lambda) \in \underset{F}{\operatorname{argmin}} & \lambda \|F\|_{(1,1)} \; + \; \|X - \phi F\|^2_{fro}, \end{cases} \tag{35}$$

where $\|M\|_{(1,1)} = \sum\limits_{i,j} |M_{ij}|$, and $\|M\|^2_{fro} = \sum\limits_{i,j} |M_{ij}|^2$.

The problems (35) are solved for a range of values of $(\epsilon_t)_t$ and $(\lambda_t)_t$, for $t = 1, 2, \ldots, T$. The solutions $F^c(\epsilon_t)$, $F^r(\lambda_t)$ are used as initial conditions when computing $F^c(\epsilon_{t+1})$, $F^r(\lambda_{t+1})$. Since computing the solutions $F^c(\epsilon)$ and $F^r(\lambda)$ is easier for larger values of ϵ and λ, we select the sequences $\sigma_w\sqrt{mN} = \epsilon_0 \geqslant \cdots \epsilon_t \geqslant \epsilon_{t+1} \geqslant \cdots$, and $25 = \lambda_0 \geqslant \lambda_t \geqslant \lambda_{t+1} \geqslant \cdots$ in decreasing order to minimize the number of iterations required. The gradient-descent step-size $(1/\beta)$ in FLIPS is chosen such that $\gamma(h) \sim 0.01$, which is achieved by selecting $\beta^+ = 0.01(\beta/\gamma)$, whereas for FISTA, the gradient-descent step-size is chosen as $(1/L)$, where $L = \lambda_{max}(\phi^a\phi)$ is the smoothness constant.

For every $t = 1, 2, \ldots, T$, we record two metrics:

- *Normalized Mean Squared Error* in reconstruction: This measures the quality of a solution F relative to the true solution F^{tr} defined as

$$NMSE(F) = 10\log_{10}\left(\frac{\left\|\widehat{F} - F^{tr}\right\|^2_{fro}}{\|F^{tr}\|^2_{fro}}\right), \text{ where } \widehat{F} = \frac{\langle X \;,\; \phi F\rangle}{\|\phi F\|^2_{fro}} F.^{5} \tag{36}$$

  In Figure 5, we plot the mappings $t \mapsto NMSE(F^c(\epsilon_t))$ and $t \mapsto NMSE(F^r(\lambda_t))$ at various SNR levels.
- *Iterations to converge*: In Figure 6a, we record the number of iterations required $k(\epsilon_t)$, $k(\lambda_t)$ to converge for both FLIPS and FISTA as a function of ϵ_t and λ_t respectively.

5. $\widehat{F}$ is a solution obtained by optimally scaling F.

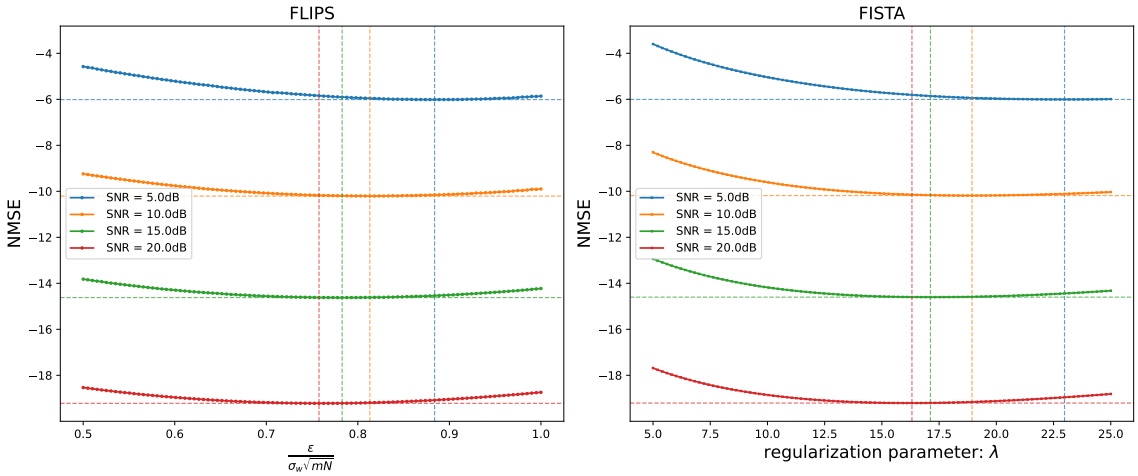


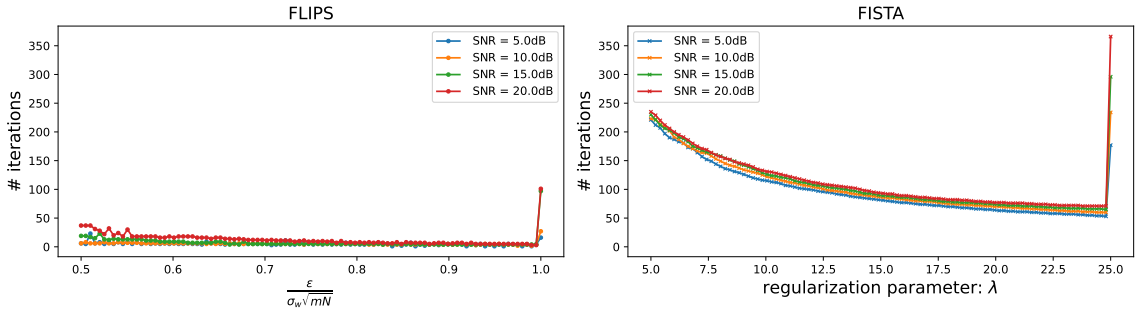
Figure 5: The Normalized Mean Squred Error (NMSE) of the optimal solution computed using FLIPS and FISTA as defined in (36). For every value of SNR, the minimum value of NMSE and the corresponding value of the parameters ϵ and λ are indicated by the intersecting horizontal and vertical dotted lines of the same color.

Convergence is defined as satisfaction of first-order optimality conditions, where the parameters for convergence criteria are tuned separately to get best result for each algorithm. Moreover, in Figure 6b, we also plot the cumulative number of iterations required $K(\epsilon_t) := \sum_{s \geqslant t} k(\epsilon_s)$, and $K(\lambda_t) := \sum_{s \geqslant t} k(\lambda_s)$ to find a solution for ϵ_t, λ_t parameterized LIP with FLIPS and FISTA respectively starting from their initial values of ϵ_0 and λ_0. In Table 3, we also report the cumulative number of iterations required to compute a solution corresponding to minimum NMSE value at a specific SNR level, i.e., the solution corresponding to the

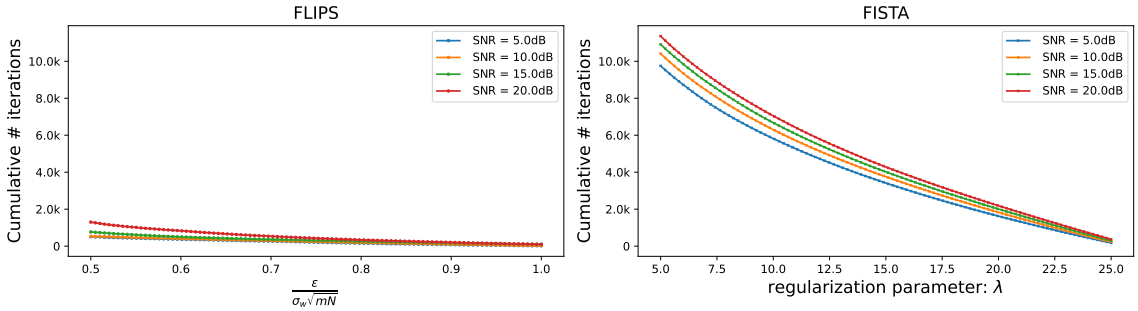
SNR levels	5dB	10dB	15dB	20dB
FLIPS (iterations)	**93.0**	**179.0**	**282.0**	**412.0**
FISTA (iterations)	731.0	2206.0	3114.0	3690.0

Table 3: Cumulative iteration required by FLIPS and FISTA to compute a minimum NMSE solution across different SNR levels.

It can be easily seen from Figure 5, that for $\epsilon \sim 0.85\sigma_w\sqrt{mN}$, the NMSE for the solution of the constrained LIP is consistently close to the minimum for a range of SNR values. Therefore, the solution obtained from FLIPS for $\epsilon = 0.85\sigma_w\sqrt{mN}$ would be satisfactory for a range of SNR levels. So, in principle, one does not have to solve the constrained LIP for a range of values of ϵ, which would significantly reduce the number of FLIPS iterations required to compute a satisfactory solution. For example, if we were to run FLIPS with $\epsilon = 0.85\sigma_w\sqrt{mN}$, the number of iterations required to converge at SNR levels: 5,10,15, and



(a) Iterations required to converge.



(b) Cumulative number of iterations required to converge.

Figure 6: Comparison of the speed of FLIPS and FISTA in solving LIPs.

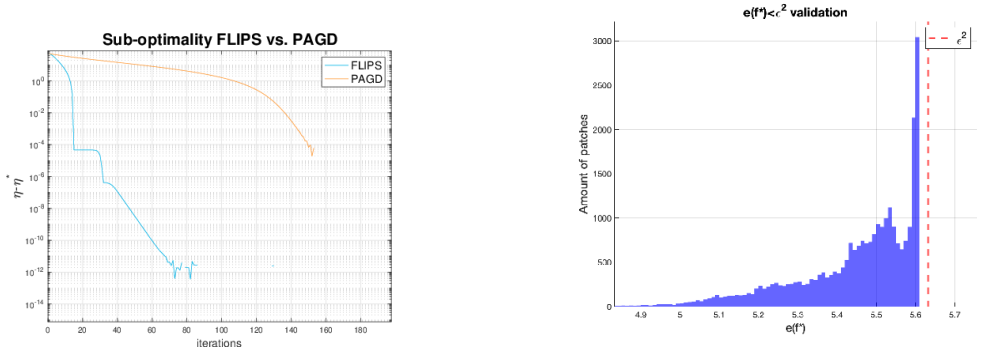
20dB would only be 29, 39, 71, and 81 respectively, which are considerably fewer than the ones reported in Table 3.

4.5 Comparison with Accelerated Projected Gradient Descent $\bar{\epsilon}$ (Remark 2.14)

In Figure 7a, we compare the sub-optimality $\eta(h_k) - c^*$ of iterates generated by FLIPS and the canonical projected accelerated gradient descent (PAGD) as in (19) (applied with $(1/b) = 2.2 \cdot 10^{-6}$ after tuning). Figure 7a clearly shows that FLIPS outperforms PAGD in terms of convergence.

4.6 Empirical validation for choosing $\bar{\epsilon}$ (Remark 2.9)

To empirically validate Remark 2.9, an experiment was conducted to check if $e(f^*) < \bar{\epsilon}^2$ by a margin. From the full 200×200 image, 28561 different 32×32 patches were extracted and the corresponding LIPs were solved to obtain the optimal solution f^* for each plot. Then the histogram of values $e(f^*)$ collected for all 28561 patches is plotted in Figure 7b (with a bandwidth of 0.01). It can be seen that there is a clear gap between the maximum $e(f^*)$ and $\bar{\epsilon}^2$. Thus, verifying empirically Remark 2.9.



(a) FLIPS vs PAGD

(b) Validation of Remark 2.9.

Figure 7: Comparison of FLIPS with projected accelerated GD (a), and in (b) the histogram of the values of $e(f^*)$ for all 32×32 patches of the 'cameraman image'.

5. Technical Proofs

Let $\Lambda := \{\lambda \in \mathbb{H} : \langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\| > 0\}$ and recall that $l(\lambda) = \sqrt{\langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\|}$.

Proposition 5.1 *Let $x, \phi,$ and ϵ be such that Assumption 2.2 holds. For any $h \in B_c$, considering the maximization problem*

$$\begin{cases} \sup \limits_{\lambda \in \Lambda} & L\left(\lambda, h\right) := 2l(\lambda) \; - \; \langle \lambda \, , \, \phi(h) \rangle \, , \end{cases} \tag{37}$$

the following assertions hold.

(i) *The maximization problem (37) is bounded if and only if $h \in \mathcal{K}(\epsilon)$, and the maximal value is equal to $\eta\left(h\right)$. In other words,*

$$\eta\left(h\right) \; = \; \sup_{\lambda \in \Lambda} \; L\left(\lambda, h\right) \qquad \textit{for all } h \in \mathcal{K}(\epsilon). \tag{38}$$

(ii) *The maximization problem (37) admits a unique maximizer $\lambda(h)$ if and only if $h \in \operatorname{int}\left(\mathcal{K}(\epsilon)\right) = \{h \in \mathbb{H} : \langle x \, , \, \phi(h) \rangle > 0, \; \textit{and } e\left(h\right) < \epsilon^2\}$, which is given by*

$$\lambda(h) = \frac{\eta\left(h\right)}{\|\phi\left(h\right)\| \sqrt{\epsilon^2 - e\left(h\right)}} \big(x - \eta\left(h\right) \phi(h)\big). \tag{39}$$

The proof of Proposition 5.1 relies heavily on (Sheriff and Chatterjee, 2020, Lemma 35, 36) under the setting $r = 2, q = 0.5$, and $\delta = 0$. We shall first provide three lemmas from which Proposition 5.1 follows easily.

Lemma 5.2 (Unboundedness in (37)) *The maximization problem (37) is unbounded for $h \notin \mathcal{K}(\epsilon)$.*

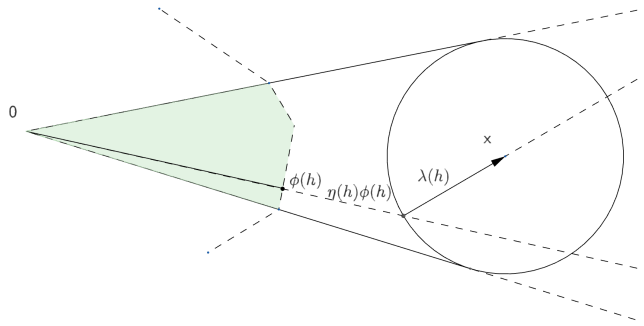


Figure 8: Graphical overview of the direction of $\lambda(h)$

Proof [Lemma 5.2] We first recall from (Sheriff and Chatterjee, 2020, Lemma 36 and assertion (iii) of Lemma 35) that the maximal value of (37) is unbounded if and only if there exists a λ' such that the two following inequalities are satisfied simultaneously

$$\left\langle \lambda' \,,\, \phi(h) \right\rangle \;\leqslant\; 0 \;<\; \left\langle \lambda' \,,\, x \right\rangle - \epsilon \left\| \lambda' \right\|. \tag{40}$$

Since $h \notin \mathcal{K}(\epsilon)$, either $\langle x \,,\, \phi\,(h) \rangle < 0$ or $e\,(h) > \epsilon^2$. On the one hand, if $\langle x \,,\, \phi\,(h) \rangle < 0$, then we observe that $\lambda' = x$ satisfies the two inequalities of (40) since $\|x\| > \epsilon$. On the other hand, if $e\,(h) > \epsilon^2$, then by considering $\lambda' = x - \frac{\langle x \,,\, \phi(h) \rangle}{\|\phi(h)\|^2} \phi(h)$, we first observe that $\langle \lambda' \,,\, \phi(h) \rangle = 0$, and by Pythagoras theorem, we have $\|\lambda'\|^2 = e\,(h)$. It is now easily verified that λ' satisfies the two inequalities (40) simultaneously since

$$\begin{cases} \left\langle \lambda' \,,\, \phi(h) \right\rangle \;=\; \langle x \,,\, \phi\,(h) \rangle - \dfrac{\langle x \,,\, \phi\,(h) \rangle}{\|\phi\,(h)\|^2} \langle \phi(h) \,,\, \phi(h) \rangle \;=\; 0 \\ \left\langle \lambda' \,,\, x \right\rangle - \epsilon \left\| \lambda' \right\| \;=\; \|x\|^2 - \dfrac{|\langle x \,,\, \phi\,(h) \rangle|^2}{\|\phi\,(h)\|^2} - \epsilon \sqrt{e\,(h)} \;=\; \sqrt{e\,(h)} \big(\sqrt{e\,(h)} - \epsilon \big) \;>\; 0. \end{cases}$$

Thus the lemma holds. ■

Lemma 5.3 (Optimal value of (37)**)** *If $h \in \mathcal{K}(\epsilon)$, then the maximal value of* (37) *is finite and equal to $\eta\,(h)$ (as in* (8)*).*

Proof [Lemma 5.3] We now recall from (Sheriff and Chatterjee, 2020, Lemma 36, and (51)-Lemma 35), that the maximal value of (37) is bounded if and only if the following minimum exists

$$\min \; \big\{ \theta \geqslant 0 : \|x - \theta \phi(h)\| \leqslant \epsilon \big\}. \tag{41}$$

Clearly, the minimum in (41) exists whenever the minimization problem is feasible. Suppose there exists some $\theta' \geqslant 0$ such that $\|x - \theta'\phi(h)\| \leqslant \epsilon$, it is immediately seen that

$$\begin{cases} \langle x\ ,\ \phi\,(h)\rangle \ \geqslant\ \dfrac{1}{2\theta'}\Big(\big(\,\|x\|^2 - \epsilon^2\big) + \theta'^2\,\|\phi\,(h)\|^2\Big) \ >\ 0, \text{ and} \\ \qquad e\,(h) \ =\ \min\limits_{\theta\in\mathbb{R}}\ \|x - \theta\phi(h)\|^2 \ \leqslant\ \big\|x - \theta'\phi(h)\big\|^2 \leqslant \epsilon^2. \end{cases}$$

Thus, $h \in \mathcal{K}(\epsilon)$. On the contrary, if $h \in \mathcal{K}(\epsilon)$, then it also seen similarly that $\theta' = \frac{\langle x\ ,\ \phi(h)\rangle}{\|\phi(h)\|^2}$ is feasible for (41). Thus, the maximal value of (37), and the minimum in (41) is finite if and only if $h \in \mathcal{K}(\epsilon)$.

It is immediately realised that the value of the minimum in (41) corresponds to the smaller root of the quadratic equation $\|x - \theta\phi(h)\|^2 = \epsilon^2$. Dividing throughout therein by θ^2, we obtain a different quadratic equation

$$\frac{1}{\theta^2}\big(\,\|x\|^2 - \epsilon^2\big) \ -\ \frac{2}{\theta}\,\langle x\ ,\ \phi\,(h)\rangle \ +\ \|\phi\,(h)\|^2 = 0.$$

Selecting the larger root (and hence smaller θ) gives us that for every $h \in \mathcal{K}(\epsilon)$, the optimal value of (37) (and (41)) is

$$\frac{\|x\|^2 - \epsilon^2}{\langle x\ ,\ \phi\,(h)\rangle \ +\ \|\phi\,(h)\|\,\sqrt{\epsilon^2 - e\,(h)}} = \eta\,(h)\,.$$

Alternatively, if one selects the smaller root of the quadratic equation $\|x - \theta\phi(h)\|^2 = \epsilon^2$, one gets the expression of eta provided in (43). ■

Remark 5.4 (Quadratic equation for $\eta\,(h)$) It is apparent from the proof of the Lemma 5.3 that for any $h \in \mathcal{K}(\epsilon)$, $\eta\,(h)$ satisfies

$$\|x - \eta\,(h)\,\phi(h)\| = \epsilon. \tag{42}$$

For $h \in \mathcal{K}(\epsilon)$, since $\langle x\ ,\ \phi\,(h)\rangle \geqslant 0$, we see that the quadratic equation $\|x - \theta\phi(h)\|^2 = \epsilon^2$ has two positive real roots. Moreover, $\eta\,(h)$ is the smallest positive root of this quadratic equation, which gives

$$\eta\,(h) = \frac{\langle x\ ,\ \phi\,(h)\rangle \ -\ \|\phi\,(h)\|\,\sqrt{\epsilon^2 - e\,(h)}}{\|\phi\,(h)\|^2} \quad \text{ for every } h \in \mathcal{K}(\epsilon). \tag{43}$$

Lemma 5.5 *Suppose, $\lambda(h)$ is an optimal solution to the maximization problem* (37)*, then it also satisfies*

$$\eta\,(h) \ =\ L\,(\lambda(h), h) \ =\ \sqrt{\langle\lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|}. \tag{44}$$

Proof [Lemma 5.5] First of all, we observe that the maximization problem (37) admits an optimal solution if and only if it satisfies the first order optimality conditions:

$$0 \ =\ \frac{\partial}{\partial\lambda}L\,(\lambda(h), h) \ =\ \frac{x - \frac{\epsilon}{\|\lambda(h)\|}\lambda(h)}{\sqrt{\langle\lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|}} \ -\ \phi(h).$$

By taking inner product throughout with $\lambda(h)$, it is readily seen that

$$\sqrt{\langle \lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|}\ =\ \langle \lambda(h)\ ,\ \phi(h)\rangle\,.$$

Thus, we have

$$\begin{aligned}\eta\,(h) &= L\,(\lambda(h), h) \quad \text{from Lemma 5.3,}\\ &= 2\sqrt{\langle \lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|} - \langle \lambda(h)\ ,\ \phi(h)\rangle\\ &= \sqrt{\langle \lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|}.\end{aligned}$$

■

Proof [Proof of Proposition 5.1] Lemma 5.2 and 5.3 together imply assertion (i) of the proposition. To complete the proof of the proposition, it now only remains to be shown that the maximization problem (37) admits a unique optimal solution $\lambda(h)$ if and only if $h \in \text{int}(\mathcal{K}(\epsilon))$.

Firstly, we observe that the maximization problem (37) admits an optimal solution if and only if it satisfies the first order optimality conditions: $0 = \frac{\partial}{\partial \lambda} L\,(\lambda(h), h)$, rearranging terms, we obtain

$$\frac{\epsilon}{\|\lambda(h)\|}\lambda(h)\ =\ x\ -\ \sqrt{\langle \lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|}\,\phi(h)\ =\ x\ -\ \eta\,(h)\,\phi(h),^{6} \tag{45}$$

the last equality is due to (44). Now, we observe that any $\lambda(h)$ that satisfies the implicit non-linear equation (45) must be of the form $\lambda(h) = r\big(x - \eta\,(h)\,\phi(h)\big)$ for some $r > 0$. The precise value of $r > 0$ can be computed using (44). We have

$$\begin{aligned}\eta\,(h) &= \sqrt{\langle \lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|}\\ &= \sqrt{r}\sqrt{\langle x - \eta\,(h)\,\phi(h)\ ,\ x\rangle - \epsilon\,\|x - \eta\,(h)\,\phi(h)\|}\\ &= \sqrt{r}\sqrt{\|x - \eta\,(h)\,\phi(h)\|^2 + \langle x - \eta\,(h)\,\phi(h)\ ,\ \eta\,(h)\,\phi(h)\rangle - \epsilon^2}\\ &= \sqrt{r\eta\,(h)}\sqrt{\langle x\ ,\ \phi\,(h)\rangle - \eta\,(h)\,\|\phi\,(h)\|^2},\end{aligned}$$

from which it is easily picked that $r\ =\ \langle x\ ,\ \phi\,(h)\rangle - \eta\,(h)\,\|\phi\,(h)\|^2\ =\ \|\phi\,(h)\|\,\sqrt{\epsilon^2 - e\,(h)}$. Now, $r > 0$ if and only if $e\,(h) < \epsilon^2$, or equivalently, $h \in \text{int}(\mathcal{K}(\epsilon))$. Thus, we finally conclude that the optimality condition (45) has a unique solution $\lambda(h)$ if and only if $h \in \text{int}(\mathcal{K}(\epsilon))$, and is given by

$$\lambda(h) = \frac{\eta\,(h)}{\|\phi\,(h)\|\,\sqrt{\epsilon^2 - e\,(h)}}\big(x - \eta\,(h)\,\phi(h)\big).$$

The proof of the proposition is complete. ■

6. Observe that by evaluating squared norm on both sides of (45) and using (44) also gives rise to the quadratic equation $\epsilon^2 = \|x - \eta\,(h)\,\phi(h)\|^2$ for $\eta\,(h)$.

Lemma 5.6 *For every $h \in \mathcal{K}(\epsilon)$, the following relations hold*

$$\eta\,(h)\,\|\phi\,(h)\| \;\geqslant\; \|x\| - \epsilon, \tag{46}$$

$$\langle x - \eta\,(h)\,\phi(h)\;,\;\phi(h)\rangle \;=\; \|\phi\,(h)\|\,\sqrt{\epsilon^2 - e\,(h)}. \tag{47}$$

Proof For any $h \in \mathcal{K}(\epsilon)$, we see from (8) that

$$\begin{aligned}\frac{1}{\eta\,(h)} &= \frac{\|\phi\,(h)\|}{\big(\,\|x\|^2 - \epsilon^2\big)}\Big(\frac{\langle x\;,\;\phi\,(h)\rangle}{\|\phi\,(h)\|} + \sqrt{\epsilon^2 - e\,(h)}\Big)\\ &< \frac{\|\phi\,(h)\|}{\big(\,\|x\|^2 - \epsilon^2\big)}\big(\,\|x\| + \epsilon\big) \quad \text{due to C-S inequality, and } 0 \leqslant e\,(h)\,,\\ &= \frac{\|\phi\,(h)\|}{\|x\| - \epsilon},\end{aligned}$$

On rearranging terms, the inequality (46) is obtained at once. Similarly, rearranging terms in (42), we see

$$\|\phi\,(h)\|\,\sqrt{\epsilon^2 - e\,(h)} \;=\; \langle x\;,\;\phi\,(h)\rangle - \eta\,(h)\,\|\phi\,(h)\|^2\,,$$

Observing that $\langle x\;,\;\phi\,(h)\rangle - \eta\,(h)\,\|\phi\,(h)\|^2 = \langle x - \eta\,(h)\,\phi(h)\;,\;\phi(h)\rangle$, (47) follows immediately. ∎

Lemma 5.7 (Derivative of $\lambda(h)$) *Let x, ϕ, ϵ be given such that Assumption 2.2 holds, then the mapping $\mathrm{int}(\mathcal{K}(\epsilon)) \ni h \longmapsto \lambda(h)$ is continuously differentiable. Moreover, with the continuous maps $\mathrm{int}(\mathcal{K}(\epsilon)) \ni h \longmapsto \big(r(h),\; M(h)\big) \in (0, +\infty) \times \mathbb{R}^{n\times n}$ defined as*

$$\begin{cases} r(h) \coloneqq \dfrac{2\epsilon}{\|\lambda(h)\|} + \dfrac{\big(\,\|x\|^2 - \epsilon^2\big)}{\eta^2(h)}, \quad \textit{and}\\ M(h) \coloneqq \dfrac{\|\lambda(h)\|}{\epsilon\eta\,(h)}\Big(\eta^2(h)\mathbb{I}_n \;+\; r(h)\big(\lambda(h)\lambda^\top(h)\big) - \big(\lambda(h)x^\top + x\lambda^\top(h)\big)\Big), \end{cases} \tag{48}$$

the derivative of $h \longmapsto \lambda(h)$ is the linear map $\left(\frac{\partial\lambda(h)}{\partial h}\right) : \mathbb{H} \longrightarrow \mathbb{R}^n$ given by

$$\left(\frac{\partial\lambda(h)}{\partial h}\right)(v) \;=\; -M(h)\cdot\phi(v) \textit{ for all } v \in \mathbb{H}. \tag{49}$$

Proof [Lemma 5.7] First, we rewrite (39) as

$$\langle x - \eta\,(h)\,\phi(h)\;,\;\phi(h)\rangle\lambda(h) \;=\; \eta\,(h)\,\big(x - \eta\,(h)\,\phi(h)\big),$$

then by differentiating on both sides w.r.t. h, we obtain an equation in the space of linear operators from $\mathbb{H}$ to $\mathbb{R}^n$. Evaluating the operators on the both sides of this equation at some $v \in \mathbb{H}$, we get

$$\begin{aligned}&\langle x - \eta\,(h)\,\phi(h)\;,\;\phi(h)\rangle\left(\frac{\partial\lambda(h)}{\partial h}\right)(v) \;+\; \Big\langle\nabla\Big(\langle x - \eta\,(h)\,\phi(h)\;,\;\phi(h)\rangle\Big)\;,\;v\Big\rangle\lambda(h)\\ &\quad= \langle\nabla\eta(h)\;,\;v\rangle x \;-\; \eta^2(h)\phi(v) \;-\; 2\eta\,(h)\langle\nabla\eta(h)\;,\;v\rangle\;\phi(h).\end{aligned} \tag{50}$$

On the one hand, we have

$$\begin{aligned}
&\langle \nabla\eta(h)\ ,\ v\rangle x \ -\ \eta^2(h)\phi(v) \ -\ 2\eta\,(h)\,\langle \nabla\eta(h)\ ,\ v\rangle\ \phi(h)\\
&\quad =\ \langle \nabla\eta(h)\ ,\ v\rangle\big(x-2\eta\,(h)\,\phi(h)\big) \ -\ \eta^2(h)\phi(v)\\
&\quad =\ -\langle \lambda(h)\ ,\ \phi(v)\rangle\big(x-2\eta\,(h)\,\phi(h)\big) \ -\ \eta^2(h)\phi(v) \quad \text{since } \nabla\eta(h)=-\phi^a(\lambda(h))\\
&\quad =\ -\Big(\eta^2(h)\,\mathbb{I}_n \ +\ \big(x-2\eta\,(h)\,\phi(h)\big)\lambda^\top(h)\Big)\cdot\phi(v) \qquad (51)\\
&\quad =\ -\Big(\eta^2(h)\,\mathbb{I}_n \ +\ \big(-x+\ {}^{2\epsilon}\!/_{\|\lambda(h)\|}\ \lambda(h)\big)\lambda^\top(h)\Big)\cdot\phi(v)\\
&\quad =\ -\Big(\eta^2(h)\,\mathbb{I}_n \ +\ {}^{2\epsilon}\!/_{\|\lambda(h)\|}\ \big(\lambda(h)\lambda^\top(h)\big) \ -\ \big(x\lambda^\top(h)\big)\Big)\cdot\phi(v).
\end{aligned}$$

On the other hand, since

$$\begin{aligned}
\nabla\Big(\langle x-\eta\,(h)\,\phi(h)\ ,\ \phi(h)\rangle\Big) \ &=\ \phi^a(x) \ -\ 2\eta\,(h)\,\phi^a\big(\phi(h)\big) \ -\ \|\phi\,(h)\|^2\,\nabla\eta(h)\\
&=\ \phi^a\Big(x-2\eta\,(h)\,\phi(h)+\|\phi\,(h)\|^2\,\lambda(h)\Big),
\end{aligned}$$

we also have

$$\begin{aligned}
&\Big\langle \nabla\Big(\langle x-\eta\,(h)\,\phi(h)\ ,\ \phi(h)\rangle\Big)\ ,\ v\Big\rangle\lambda(h)\\
&\qquad =\ \Big\langle \big(x-2\eta\,(h)\,\phi(h)+\|\phi\,(h)\|^2\,\lambda(h)\big)\ ,\ \phi(v)\Big\rangle\lambda(h)\\
&\qquad =\ \Big(\lambda(h)\big(x-2\eta\,(h)\,\phi(h)+\|\phi\,(h)\|^2\,\lambda(h)\big)^\top\Big)\cdot\phi(v) \qquad (52)\\
&\qquad =\ \Big(\lambda(h)\big(-x+\big({}^{2\epsilon}\!/_{\|\lambda(h)\|}+\|\phi\,(h)\|^2\big)\lambda(h)\big)^\top\Big)\cdot\phi(v) \quad \text{from (39)}\\
&\qquad =\ \Big(-\big(\lambda(h)x^\top\big) \ +\ \big({}^{2\epsilon}\!/_{\|\lambda(h)\|}+\|\phi\,(h)\|^2\big)\big(\lambda(h)\lambda^\top(h)\big)\Big)\cdot\phi(v).
\end{aligned}$$

Collecting (51) and (52), together with $\langle x-\eta\,(h)\,\phi(h)\ ,\ \phi(h)\rangle = {}^{\epsilon\eta(h)}\!/_{\|\lambda(h)\|}$ (from (39)), (50) simplifies to reveal

$$\begin{aligned}
\frac{\epsilon\eta\,(h)}{\|\lambda(h)\|}\left(\frac{\partial\lambda(h)}{\partial h}\right)(v) \ &=\ -\Big(\eta^2(h)\mathbb{I}_n \ -\ \big(x\lambda^\top(h) \ +\ \lambda(h^{x^\top})\big)\Big)\cdot\phi(v)\\
&\qquad -\Big({}^{4\epsilon}\!/_{\|\lambda(h)\|} \ +\ \|\phi\,(h)\|^2\Big)\big(\lambda(h)\lambda^\top(h)\big)\cdot\phi(v).
\end{aligned}$$

Finally, simplifying

$$\begin{aligned}
\frac{2\epsilon}{\|\lambda(h)\|} \ +\ \|\phi\,(h)\|^2 \ &=\ \frac{2\langle x-\eta\,(h)\,\phi(h)\ ,\ \phi(h)\rangle}{\eta\,(h)} \ +\ \|\phi\,(h)\|^2 \quad \text{from (39)}\\
&=\ \frac{2\langle x\ ,\ \phi\,(h)\rangle}{\eta\,(h)} \ -\ \|\phi\,(h)\|^2\\
&=\ \frac{1}{\eta^2(h)}\Big(2\langle x\ ,\ \eta\,(h)\,\phi(h)\rangle \ -\ \eta^2(h)\,\|\phi\,(h)\|^2\Big)\\
&=\ \frac{1}{\eta^2(h)}\Big(\|x\|^2-\|x-\eta\,(h)\,\phi(h)\|^2\Big) \ =\ \frac{\big(\|x\|^2-\epsilon^2\big)}{\eta^2(h)}.
\end{aligned}$$

Putting everything together, the derivative $\left(\frac{\partial\lambda(h)}{\partial h}\right)(v)$ is easily written in terms of the matrix $M(h)$ given in (48) as

$$\left(\frac{\partial\lambda(h)}{\partial h}\right)(v) \;=\; M(h)\cdot\phi(v)$$

Continuous differentiability of $\mathrm{int}(\mathcal{K}(\epsilon))\ni h\longmapsto\lambda(h)\in\mathbb{R}^n$ follows directly from continuity of the map $\mathrm{int}(\mathcal{K}(\epsilon))\ni h\longmapsto M(h)\in\mathbb{R}^{n\times n}$, which is straight forward. The proof of the lemma is complete. ■

Proof [Proof of Proposition 2.5] From assertion (i) of Proposition 5.1, it is inferred that for every $h\in\mathcal{K}(\epsilon)$, the value $\eta(h)$ is a point-wise maximum of the linear function $L(\lambda,h)$ (linear in h). Thus, the mapping $\eta:\mathcal{K}(\epsilon)\longrightarrow[0,+\infty)$ is convex.

From assertion (ii) of Proposition 5.1, it follows that the maximization problem (37) admits a solution $\lambda(h)$ if and only if $h\in\mathcal{K}(\bar{\epsilon})$.Then, from Danskin's theorem (Bertsekas, 1971), we conclude that the function $\eta:\mathcal{K}(\epsilon)\longrightarrow[0,+\infty)$ is differentiable if and only if the maximizer $\lambda(h)$ in (37) exists. Thus, $\eta:\mathcal{K}(\epsilon)\longrightarrow[0,+\infty)$ is differentiable at every $h\in\mathrm{int}(\mathcal{K}(\epsilon))$, and the derivative is given by $\nabla\eta(h)=-\phi^a\big(\lambda(h)\big)$. Substituting for $\lambda(h)$ from (39), we immediately get (9).

Since $\nabla\eta(h)=-\phi^a\big(\lambda(h)\big)$, we realise that $\eta(h)$ is twice differentiable if and only if the mapping $h\longmapsto\lambda(h)$ has a well-defined derivative $\left(\frac{\partial\lambda(h)}{\partial h}\right)$. In which case, the hessian is a linear operator $\big(\Delta^2\eta(h)\big):\mathbb{H}\longrightarrow\mathbb{H}$ given by

$$\big(\Delta^2\eta(h)\big)(v)=-\phi^a\circ\left(\frac{\partial\lambda(h)}{\partial h}\right)(v)\qquad\text{for all } v\in\mathbb{H}.$$

We know that the derivative $\left(\frac{\partial\lambda(h)}{\partial h}\right)$ exists for every $h\in\mathrm{int}(\mathcal{K}(\epsilon))$, thus, $\eta(\cdot)$ is twice differentiable everywhere on $\mathrm{int}(\mathcal{K}(\epsilon))$. Substituting for $\left(\frac{\partial\lambda(h)}{\partial h}\right)$ from (49), we immediately get

$$\big(\Delta^2\eta(h)\big)(v) \;=\; \left(\phi^a\circ M(h)\circ\phi\right)(v)\qquad\text{for all } v\in\mathbb{H},$$

where $h\longmapsto M(h)$ is a matrix valued map given in (48). Moreover, continuity of the hessian i.e., continuity of the map $\mathrm{int}(\mathcal{K}(\epsilon))\ni h\longmapsto\big(\Delta^2\eta(h)\big)$ follows directly from the continuity of $\mathrm{int}(\mathcal{K}(\epsilon))\ni h\longmapsto\left(\frac{\partial\lambda(h)}{\partial h}\right)$. The proof in now complete. ■

Lemma 5.8 (Smallest and largest eigenvalues of $M(h)$) *For every $h\in\mathrm{int}(\mathcal{K}(\epsilon))$, consider $M(h)\in\mathbb{R}^{n\times n}$ as given in* (48)*. Then its minimum and maximum eigenvalues, denoted by $\bar{\sigma}(M(h))$ and $\widehat{\sigma}(M(h))$ respectively, are*

$$\begin{cases}\bar{\sigma}(M(h)) \;=\; \dfrac{\left(\|x\|^2-\epsilon^2\right)\|\lambda(h)\|^3}{2\epsilon\eta^3(h)}\left(1\;-\;\sqrt{1\;-\;\dfrac{8\epsilon\eta^6(h)}{\left(\|x\|^2-\epsilon^2\right)^2\|\lambda(h)\|^3}}\right)\\[2ex] \widehat{\sigma}(M(h)) \;=\; \dfrac{\left(\|x\|^2-\epsilon^2\right)\|\lambda(h)\|^3}{2\epsilon\eta^3(h)}\left(1\;+\;\sqrt{1\;-\;\dfrac{8\epsilon\eta^6(h)}{\left(\|x\|^2-\epsilon^2\right)^2\|\lambda(h)\|^3}}\right).\end{cases}\tag{53}$$

Proof [Lemma 5.8] Recall from (48) that

$$\begin{cases} r(h) \;=\; \dfrac{2\epsilon}{\|\lambda(h)\|} + \dfrac{\big(\,\|x\|^2 - \epsilon^2\big)}{\eta^2(h)}, \quad \text{and} \\ M(h) \;=\; \dfrac{\|\lambda(h)\|}{\epsilon\eta\,(h)}\Big(\eta^2(h)\mathbb{I}_n \;+\; r(h)\big(\lambda(h)\lambda^\top(h)\big) - \big(\lambda(h)x^\top + x\lambda^\top(h)\big)\Big). \end{cases}$$

First, suppose that $\lambda(h)$ and x are linearly independent. Then it is clear that the subspace $S \coloneqq \text{span}\{\lambda(h), x\}$ is invariant under the linear transformation given by the matrix $M(h)$, and this linear transformation is identity on the orthogonal complement of S. Then it is also evident that the hessian has $n-2$ eigenvalues equal to $\big(1/\epsilon\big)\eta\,(h)\,\|\lambda(h)\|$ and the two other distinct eigenvalues corresponding to the restriction of $M(h)$ onto the 2-dimensional subspace S.

Let T denote the 2×2 matrix representing the restriction of $M(h)$ onto the subspace S for $\{\lambda(h),\ x\}$ being chosen as a basis for S. In other words, it holds that $M(h)[x \quad \lambda(h)] = [x \quad \lambda(h)]T$. Using the fact that $\eta^2(h) = \langle\lambda(h)\ ,\ x\rangle - \epsilon\,\|\lambda(h)\|$ from (44), it is easily verified that the matrix T simplifies to

$$T = \frac{\|\lambda(h)\|}{\epsilon\eta\,(h)}\begin{pmatrix} -\epsilon\,\|\lambda(h)\| & -\|\lambda(h)\|^{\,2} \\ r(h)\,\langle\lambda(h)\ ,\ x\rangle - \|x\|^2 & r(h)\,\|\lambda(h)\|^{\,2} - \epsilon\,\|\lambda(h)\| \end{pmatrix}. \tag{54}$$

Furthermore, substituting $r(h)$, it is also verified that $\text{tr}(T) = \frac{\|\lambda(h)\|^2}{\eta^2(h)}\big(\,\|x\|^2 - \epsilon^2\big)$ and $\det(T) = 2\epsilon\eta^2(h)\,\|\lambda(h)\|$. Now, it is easily verified that the two eigenvalues of T are precisely equal to $\{\bar\sigma(M(h)), \widehat\sigma(M(h))\}$.

Since apart from $\{\bar\sigma(M(h)), \widehat\sigma(M(h))\}$, the rest of the eigenvalues of $M(h)$ are equal to $\frac{1}{\epsilon}\eta\,(h)\,\|\lambda(h)\|$, it remains to be shown that $\bar\sigma(M(h)) \leqslant \frac{1}{\epsilon}\eta\,(h)\,\|\lambda(h)\| \leqslant \widehat\sigma(M(h))$; which we do so by producing $u_1, u_2 \in S$ such that

$$\bar\sigma(M(h)) \;\leqslant\; \frac{\langle u_1\ ,\ M(h)u_1\rangle}{\|u_1\|^2} \;\leqslant\; \frac{1}{\epsilon}\eta\,(h)\,\|\lambda(h)\| \;\leqslant\; \frac{\langle u_2\ ,\ M(h)u_2\rangle}{\|u_2\|^2} \;\leqslant\; \widehat\sigma(M(h)). \tag{55}$$

Observe that the inequalities $\bar\sigma(M(h)) \;\leqslant\; \frac{\langle u_1\ ,\ M(h)u_1\rangle}{\|u_1\|^2}$, and $\frac{\langle u_2\ ,\ M(h)u_2\rangle}{\|u_2\|^2} \;\leqslant\; \widehat\sigma(M(h))$ readily hold for any $u_1, u_2 \in S$ since $\bar\sigma(M(h)), \widehat\sigma(M(h))$ are the two eigenvalues of $M(h)$ when restricted to the subspace S. To obtain the rest of the inequalities in (55), consider

$$u_1 \;=\; x \;+\; \frac{\|x\|^2 - \;r(h)\,\langle\lambda(h)\ ,\ x\rangle}{r(h)\,\|\lambda(h)\|^{\,2} - \;\langle\lambda(h)\ ,\ x\rangle}\lambda(h) \quad \text{and} \quad u_2 \;=\; x \;-\; \frac{\|x\|^2}{\langle\lambda(h)\ ,\ x\rangle}\lambda(h).$$

It is easily verified that $\langle u_1\ ,\ r(h)\lambda(h) - x\rangle \;=\; 0$ and $\langle u_2\ ,\ x\rangle \;=\; 0$. Moreover, rewriting $M(h)$ by completing squares as

$$M(h) = \frac{\|\lambda(h)\|}{\epsilon r(h)\eta\,(h)}\Big(\eta^2(h)r(h)\mathbb{I}_n \;+\; \big(r(h)\lambda(h) - x\big)\big(r(h)\lambda(h) - x\big)^\top - \;xx^\top\Big),$$

it is also easily verified that the inequalities

$$\begin{cases} \dfrac{\langle u_1\ ,\ M(h)u_1\rangle}{\|u_1\|^2} \;=\; \dfrac{\|\lambda(h)\|}{\epsilon r(h)\eta\,(h)}\Big(\eta^2(h) \;-\; \dfrac{|\langle u_1\ ,\ x\rangle|^2}{\|u_1\|^2}\Big) & \leqslant\; \dfrac{1}{\epsilon}r(h)\,\|\lambda(h)\|\,, \\ \dfrac{\langle u_2\ ,\ M(h)u_2\rangle}{\|u_2\|^2} \;=\; \dfrac{\|\lambda(h)\|}{\epsilon r(h)\eta\,(h)}\Big(\eta^2(h) \;+\; \dfrac{|\langle u_2\ ,\ r(h)\lambda(h) - x\rangle|^2}{\|u_2\|^2}\Big) & \geqslant\; \dfrac{1}{\epsilon}r(h)\,\|\lambda(h)\|\,. \end{cases}$$

Thus, the inequalities (55) are obtained at once.

To complete the proof for the case when $\lambda(h)$ and x are linearly dependent, we first see that the $\text{int}(\mathcal{K}(\epsilon)) \ni h \longmapsto \big(\bar{\sigma}(M(h)), \widehat{\sigma}(M(h))\big)$ is continuous. Secondly, since the mapping $\text{int}(\mathcal{K}(\epsilon)) \ni h \longmapsto M(h)$ is also continuous, and the eigenvalues of a matrix vary continuously, these two limits must be the same. The proof is now complete. ■

Proof [Proposition 2.7] For any $\bar{\epsilon} \in (0, \epsilon)$ and $\widehat{\eta} > c^*$, we know that the set $\mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \subset \text{int}(\mathcal{K}(\epsilon))$. Consequently, it follows from Proposition 2.5 that $\eta : \mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \longrightarrow [0, +\infty)$ is twice continuously differentiable. To establish the required smoothness, and strong convexity assertions of the proposition, we first obtain uniform upper (lower) bound on the maximum (minimum) eigenvalue of the Hessian $\big(\Delta^2\eta(h)\big)$. To this end, for every $v \in \mathbb{H}$, since $\big\langle v\ ,\ \big(\Delta^2\eta(h)\big)(v)\big\rangle = \langle \phi(v)\ ,\ M(h)\phi(v)\rangle$, we see that

$$\bar{\sigma}(M(h))\bar{\sigma}(\phi^a \circ \phi) \ \leqslant\ \frac{\big\langle v\ ,\ \big(\Delta^2\eta(h)\big)(v)\big\rangle}{\|v\|^2} \ \leqslant\ \widehat{\sigma}(M(h))\widehat{\sigma}(\phi^a \circ \phi) \quad \text{for all } v \in \mathbb{H}. \tag{56}$$

The quantities $\bar{\sigma}(\phi^a \circ \phi)$ and $\widehat{\sigma}(\phi^a \circ \phi)$ are the minimum and maximum eigenvalues of the linear operator $\phi^a \circ \phi : \mathbb{H} \longrightarrow \mathbb{H}$ respectively. Denoting $\widehat{\sigma}\big(\Delta^2\eta(h)\big)$ and $\widehat{\sigma}\big(\Delta^2\eta(h)\big)$ to be the the maximum and minimum eigenvalues of the hessian respectively, it follows from (56) that

$$\bar{\sigma}(M(h))\bar{\sigma}(\phi^a \circ \phi) \ \leqslant\ \bar{\sigma}\big(\Delta^2\eta(h)\big) \ \leqslant\ \widehat{\sigma}\big(\Delta^2\eta(h)\big) \ \leqslant\ \widehat{\sigma}(M(h))\widehat{\sigma}(\phi^a \circ \phi). \tag{57}$$

Uniform upper bound for $\widehat{\sigma}(M(h))$. For every $h \in \mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \subset \mathcal{H}$, we have the inequality

$$\begin{aligned} \frac{\|\lambda(h)\|}{\eta(h)} \ &=\ \frac{\|x - \eta(h)\,\phi(h)\|}{\|\phi(h)\|\sqrt{\epsilon^2 - e(h)}} \ =\ \frac{\epsilon}{\|\phi(h)\|\sqrt{\epsilon^2 - e(h)}},\ \text{from (42)}, \\ &<\ \frac{\epsilon}{\|\phi(h)\|\sqrt{\epsilon^2 - \bar{\epsilon}^2}},\ \text{since } e(h) < \bar{\epsilon}^2 \text{ for } h \in \mathcal{H}(\bar{\epsilon}, \widehat{\eta}). \end{aligned} \tag{58}$$

On the other hand, since $\mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \subset \mathcal{K}(\epsilon)$ we conclude from (46) that the upper bound $\frac{1}{\|\phi(h)\|} < \frac{\eta(h)}{\|x\| - \epsilon}$ holds for every $h \in \mathcal{H}(\bar{\epsilon}, \widehat{\eta})$. Putting together in (58), we have

$$\frac{\|\lambda(h)\|}{\eta(h)} \ <\ \frac{\eta(h)}{\|x\| - \epsilon}\frac{\epsilon}{\sqrt{\epsilon^2 - \bar{\epsilon}^2}} \ <\ \frac{\widehat{\eta}}{\|x\| - \epsilon}\frac{\epsilon}{\sqrt{\epsilon^2 - \bar{\epsilon}^2}}\ .$$

Thus, from (53), we have

$$\widehat{\sigma}(M(h)) \ \leqslant\ \frac{1}{\epsilon}\big(\|x\|^2 - \epsilon^2\big)\left(\frac{\|\lambda(h)\|}{\eta(h)}\right)^3 \ <\ \frac{\epsilon^2\big(\|x\| + \epsilon\big)}{\big(\|x\| - \epsilon\big)^2}\left(\frac{\widehat{\eta}}{\sqrt{\epsilon^2 - \bar{\epsilon}^2}}\right)^3. \tag{59}$$

Uniform lower bound for $\bar{\sigma}(M(h))$. We know that $\sqrt{1 - \theta^2} < 1 - \frac{\theta^2}{2}$ for every $\theta \in [0, 1]$. Using this inequality in (53) for $\bar{\sigma}(M(h))$ and simplifying, we see that

$$\bar{\sigma}(M(h)) \ \geqslant\ \frac{2\eta^3(h)}{\big(\|x\|^2 - \epsilon^2\big)} \ \geqslant\ \frac{2c^{*3}}{\big(\|x\|^2 - \epsilon^2\big)} \geqslant\ \frac{2\bar{\eta}^3}{\big(\|x\|^2 - \epsilon^2\big)}, \tag{60}$$

for every $\bar{\eta} \in (0, c^*]$. Collecting (59) and (60), we see that the minimum and maximum eigenvalues of the hessian are uniformly bounded over $\mathcal{H}(\bar{\epsilon}, \widehat{\eta})$, and the bounds are

$$\begin{cases} \widehat{\sigma}\left(\Delta^2 \eta(h)\right) & < \dfrac{\epsilon^2 \widehat{\eta}^3}{\left(\epsilon^2 - \bar{\epsilon}^2\right)^{3/2}} \dfrac{\left(\|x\| + \epsilon\right)}{\left(\|x\| - \epsilon\right)^2} \widehat{\sigma}(\phi^a \circ \phi) & =: \beta(\bar{\epsilon}, \widehat{\eta}), \\ \bar{\sigma}\left(\Delta^2 \eta(h)\right) & \geqslant \dfrac{2\bar{\eta}^3}{\left(\|x\|^2 - \epsilon^2\right)} \bar{\sigma}(\phi^a \circ \phi) & =: \alpha(\bar{\eta}). \end{cases} \tag{61}$$

Finally, $\eta : \mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \longrightarrow [0, +\infty)$ is twice continuously differentiable with the maximum eigenvalue of the hessian being uniformly bounded above by $\beta(\bar{\epsilon}, \widehat{\eta})$. It then follows that $\eta : \mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \longrightarrow [0, +\infty)$ is $\beta(\bar{\epsilon}, \widehat{\eta})$-smooth in the sense of (11). Moreover, if ϕ is invertible in addition, then the minimum eigenvalue of the hessian is uniformly bounded below by $\alpha > 0$. Consequently, the mapping $\eta : \mathcal{H}(\bar{\epsilon}, \widehat{\eta}) \longrightarrow [0, +\infty)$ is α-strongly convex in the sense of (12). The proof of the proposition is now complete. ■

5.1 Proofs for reformulation as a smooth minimization problem

Lemma 5.9 (Non-smooth reformulation) *Consider the LIP* (1) *under the setting of Assumption 2.2, then the LIP* (1) *is equivalent to the minimization problem*

$$\begin{cases} \min\limits_{h \in B_c \cap \mathcal{K}(\epsilon)} \eta(h) \quad . \end{cases} \tag{62}$$

In other words, the optimal value of (14) *is equal to* c^* *and* h^* *is a solution to* (14) *if and only if* $c^* h^*$ *is an optimal solution to* (1).

Proof [Lemma 5.9] Recall that $\Lambda = \{\lambda \in \mathbb{R}^n : \langle \lambda\ ,\ x\rangle - \epsilon \|\lambda\| > 0\}$, $B_c = \{h \in \mathbb{H} : c(\mathbb{H}) \leqslant 1\}$, and $L(\lambda, h) = 2\sqrt{\langle \lambda\ ,\ x\rangle - \epsilon \|\lambda\|} - \langle \lambda\ ,\ \phi(h)\rangle$. The original LIP (1) was reformulated as the min-max problem. By considering $r = 2, q = 0.5, \delta = 0$ in (Sheriff and Chatterjee, 2020, Theorem 10) we see that the min-max problem

$$\begin{cases} \min\limits_{h\ \in\ B_c}\ \sup\limits_{\lambda\ \in\ \Lambda}\ L(\lambda, h), \end{cases} \tag{63}$$

is equivalent to the LIP (1) with the optimal value of the min-max problem equal to c^*. Moreover, from (Sheriff and Chatterjee, 2020, Theorem 10, assertion (ii)-a), it also follows that $h^* \in \underset{h\ \in\ B_c}{\operatorname{argmin}} \left\{ \sup\limits_{\lambda\ \in\ \Lambda} L(\lambda, h) \right\}$ if and only if $c^* h^*$ is an optimal solution to the LIP (1). Solving for the maximization problem over λ in the min-max problem (63), in view of Proposition 5.1 we know that the maximum over λ is equal to $\eta(h)$ whenevr it is finite. Therefore, we get

$$h^* \in \underset{h \in B_c \cap \mathcal{K}(\epsilon)}{\operatorname{argmin}} \eta(h),$$

if and only if $c^* h^*$ is an optimal solution to the LIP (1). The proof is now complete. ■

Proof [Theorem 2.8] Under the setting of Assumption 2.2 we have $B(x,\epsilon) \cap \text{image}(\phi) \neq \varnothing$. Thus, it follows from (Sheriff and Chatterjee, 2020, Proposition 31-(ii)) and consequently, from (Sheriff and Chatterjee, 2020, Theorem 10-(ii)-b), that the min-max problem

$$\begin{cases} \min\limits_{h \in B_c} \sup\limits_{\lambda \in \Lambda} & 2\sqrt{\langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\|} \; - \; \langle \lambda \, , \, \phi(h) \rangle \end{cases}$$

admits a saddle point solution. Moreover, every saddle point $(h^*, \lambda^*) \in B_c \times \Lambda$ is such that $h^* = \left(1/c^*\right) f^*$ where f^* is any optimal solution to the LIP (1), and λ^* is unique that satisfies

$$\lambda^* = \underset{\lambda \in \Lambda}{\text{argmax}} \quad 2\sqrt{\langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\|} \; - \; \langle \lambda \, , \, \phi(h^*) \rangle \, .$$

In view of Proposition 5.1-(ii), we conclude that $h^* \in \text{int}(\mathcal{K}(\epsilon))$. Thus, $e(f^*) = e(h^*) < \epsilon^2$, this establishes assertion (i) of the lemma.

To prove the rest of the theorem, consider any $\bar{\epsilon}, \widehat{\eta} > 0$ such that $e(f^*) \leqslant \bar{\epsilon}^2 < \epsilon^2$ and $c^* \leqslant \widehat{\eta}$. Then for any $h^* \in \underset{h \, \in \, B_c \cap \mathcal{K}(\epsilon)}{\text{argmin}} \; \eta \, (h)$, we conclude from Lemma 5.9 that $c^* h^*$ is an optimal solution to the LIP (1). Consequently, assertion (i) of the proposition then implies that $e(h^*) = e(c^* h^*) \leqslant \bar{\epsilon}^2$. Thus, we have $h^* \in \mathcal{K}(\bar{\epsilon})$. Moreover, from Lemma 5.9 it is also immediate that $\eta(h^*) = c^* \leqslant \widehat{\eta}$. Thus, $h^* \in \mathcal{H} \, (\bar{\epsilon}, \widehat{\eta})$, and we have the inclusion

$$\underset{h \, \in \, B_c \cap \mathcal{K}(\epsilon)}{\text{argmin}} \; \eta \, (h) \; \subset \; \mathcal{H} \, (\bar{\epsilon}, \widehat{\eta}) \, .$$

Since $\mathcal{H} \, (\bar{\epsilon}, \widehat{\eta}) \subset B_c \cap \mathcal{K}(\epsilon)$ to begin with, we conclude

$$\underset{h \, \in \, B_c \cap \mathcal{K}(\epsilon)}{\text{argmin}} \; \eta \, (h) \; = \; \underset{h \, \in \, \mathcal{H}(\bar{\epsilon}, \widehat{\eta})}{\text{argmin}} \; \eta \, (h) \, .$$

Now assertion (ii) of the theorem follows immediately as a consequence of Lemma 5.9. ■

5.2 Proofs for reformulation as a strongly-convex min-max problem

Proof [Lemma 2.15] Recall that $\Lambda \ni \lambda \longmapsto l(\lambda) = \sqrt{\langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\|}$, then denoting $\Lambda := \{\lambda \in \mathbb{R}^n : \langle \lambda \, , \, x \rangle - \epsilon \, \|\lambda\| > 0\}$, it is known from (Sheriff and Chatterjee, 2020) that the LIP (1) is equivalent to the min-max problem

$$\begin{cases} \min\limits_{h \in B_c} \sup\limits_{\lambda \in \Lambda} & L \, (\lambda, h) \; = \; 2l(\lambda) \; - \; \langle \lambda \, , \, \phi(h) \rangle \, . \end{cases} \tag{64}$$

In particular, under the setting of Assumption 2.2, it follows that the min-max problem (64) admits a saddle point solution. It follows from (Sheriff and Chatterjee, 2020, Theorem 10)) that a pair (h^*, λ^*) is a saddle point of (64) if and only if $c^* h^*$ is an optimal solution to the LIP (1), and $\lambda^* = \lambda(h^*)$ in view of Lemma 5.1.[7]

7. The inclusion $\lambda^* \in c^* \Lambda$ provided in (Sheriff and Chatterjee, 2020, (44), Theorem 10) turns out to be same as the condition $\lambda^* = \lambda(h^*)$ under the setting of Assumption 2.2 for LIP (1). This can be formally

We prove the lemma by establishing that every saddle point solution to the min-max problem (64) is indeed a saddle point solution to the min-max problem (21) as well. We observe that the only difference between the min-max problems (21) and (64) is in their respective feasible sets $\Lambda(\bar{\eta}, B)$ and Λ for the variable λ. Moreover, since $\Lambda(\bar{\eta}, B) \subset \Lambda$, it suffices to show that for every saddle point (h^*, λ^*) of (64), the inclusion $\lambda^* \in \Lambda(\bar{\eta}, B)$ also holds. To establish this inclusion, we first recall from (44) that

$$l(\lambda^*) = l(\lambda(h^*)) = \eta(h^*) = c^* \ \geqslant \ \bar{\eta}.$$

Secondly, using (39) we also have

$$\begin{aligned}
\|\lambda(h^*)\| \ &= \ \frac{\epsilon\, \eta(h^*)}{\|\phi(h^*)\| \sqrt{\epsilon^2 - e(h^*)}} \ \leqslant \ \frac{\epsilon\, \eta(h^*)}{\|\phi(h^*)\| \sqrt{\epsilon^2 - \bar{\epsilon}^2}} \quad \text{since } e(h^*) \in (\bar{\epsilon}^2, \epsilon^2), \\
&\leqslant \ \frac{\epsilon\, \eta^2(h^*)}{\big(\, \|x\| - \epsilon\big)\sqrt{\epsilon^2 - \bar{\epsilon}^2}} \quad \text{from (46)}, \\
&\leqslant \ \frac{\epsilon\, \bar{\eta}^2}{\big(\, \|x\| - \epsilon\big)\sqrt{\epsilon^2 - \bar{\epsilon}^2}} \ = \ B.
\end{aligned}$$

Thus, $\lambda^* \in \Lambda(\bar{\eta}, B)$ and the lemma holds. ■

Lemma 5.10 *Consider $x \in \mathbb{R}^n$ and $\epsilon > 0$ such that $\|x\| > \epsilon$. Then the following assertions hold with regards to the mapping $\Lambda \ni \lambda \longmapsto l(\lambda) := \sqrt{\langle \lambda \ , \ x\rangle - \epsilon\, \|\lambda\|}$.*

(i) *the mapping $\Lambda \ni \lambda \longmapsto l(\lambda)$ is twice continuously differentiable and its hessian $H(\lambda)$ evaluated at $\lambda \in \Lambda$ is given by*

$$H(\lambda) = \frac{-\epsilon}{2l(\lambda)\, \|\lambda\|} \left(\mathbb{I}_n \ - \ \frac{1}{\|\lambda\|^2} \lambda\lambda^\top \right) \ - \ \frac{1}{4(l(\lambda))^3} \left(x - \frac{\epsilon}{\|\lambda\|}\lambda \right) \left(x - \frac{\epsilon}{\|\lambda\|}\lambda \right)^\top . \tag{65}$$

(ii) *The smallest and largest absolute values of the eigenvalues of $H(\lambda)$ denoted respectively by $\bar{\sigma}(H(\lambda))$ and $\widehat{\sigma}(H(\lambda))$, are given by*

$$\begin{cases}
\bar{\sigma} \ = \ \dfrac{\big(\, \|x\|^2 - \epsilon^2\big)}{8(l(\lambda))^3} \left(1 \ - \ \sqrt{1 \ - \ \dfrac{8\epsilon (l(\lambda))^6}{\big(\, \|x\|^2 - \epsilon^2\big)^2\, \|\lambda\|^3}} \right) \\
\widehat{\sigma} \ = \ \dfrac{\big(\, \|x\|^2 - \epsilon^2\big)}{8(l(\lambda))^3} \left(1 \ + \ \sqrt{1 \ - \ \dfrac{8\epsilon (l(\lambda))^6}{\big(\, \|x\|^2 - \epsilon^2\big)^2\, \|\lambda\|^3}} \right).
\end{cases} \tag{66}$$

established by observing from (Sheriff and Chatterjee, 2020, Proposition 31) that $\lambda^* = c^* \frac{x - c^*\phi(h^*)}{\|x - c^*\phi(h^*)\|'_\phi}$, and then, from (Sheriff and Chatterjee, 2020, Lemma 33) we also have

$$\begin{aligned}
\big\|x - c^*\phi(h^*)\big\|'_\phi \ &= \ \max_{h \in B_c} \big\langle x - c^*\phi(h^*) \ , \ \phi(h) \big\rangle \ = \ \big\langle x - c^*\phi(h^*) \ , \ \phi(h^*) \big\rangle \\
&= \ \big\|\phi(h^*)\big\| \sqrt{\epsilon^2 - e(h^*)}.
\end{aligned}$$

Proof [Lemma 5.10] First of all, we observe that since $\lambda \longmapsto l(\lambda)$ is differentiable everywhere on Λ, and the gradients are given by $\nabla l(\lambda) = \frac{1}{2l(\lambda)}\Big(x - \frac{\epsilon}{\|\lambda\|}\lambda\Big)$. Differentiating again w.r.t. λ, we easily verify that the hessian is indeed as given by (65). First, suppose that λ and x are linearly independent, observe that the subspace $S := \text{span}\{\lambda, x - \frac{\epsilon}{\|\lambda\|}\lambda\}$ is invariant under the linear transformation given by the hessian matrix $H(\lambda)$, and it is identity on the orthogonal complement of S. Then it is evident that the hessian has $n-2$ eigenvalues equal to $-\epsilon/l(\lambda)\|\lambda\|$ and the two other distinct eigenvalues corresponding to the restriction of $H(\lambda)$ onto S. Selecting $\{\lambda,\ x - \frac{\epsilon}{\|\lambda\|}\lambda\}$ as a basis for S, the linear mapping of the hessian is given by the matrix

$$T = \begin{pmatrix} 0 & \frac{\epsilon l(\lambda)}{2\|\lambda\|^3} \\ \frac{-1}{4l(\lambda)} & \frac{-\left(\|x\|^2 - \epsilon^2\right)}{4(l(\lambda))^3} \end{pmatrix}. \tag{67}$$

It is a straightforward exercise to verify that $-\bar{\sigma}$ and $-\widehat{\sigma}$ are indeed the two distinct eigenvalues of T and consequently, the remaining two eigenvalues of the hessian $H(\lambda)$. Since the rest of the eigenvalues are $-\epsilon/l(\lambda)\|\lambda\|$, it remains to be shown that $\bar{\sigma} \leqslant \epsilon/2l(\lambda)\|\lambda\| \leqslant \widehat{\sigma}$. We establish it by producing $u_1, u_2 \in S$ such that

$$\bar{\sigma} \ \leqslant\ \frac{|\langle u_1\ ,\ H(\lambda)u_1\rangle|}{\|u_1\|^2} \ \leqslant\ \frac{\epsilon}{2l(\lambda)\,\|\lambda\|} \ \leqslant\ \frac{|\langle u_2\ ,\ H(\lambda)u_2\rangle|}{\|u_2\|^2} \ \leqslant\ \widehat{\sigma}. \tag{68}$$

Observe that the inequalities $\bar{\sigma} \ \leqslant\ \frac{|\langle u_1\ ,\ H(\lambda)u_1\rangle|}{\|u_1\|^2}$, and $\frac{|\langle u_2\ ,\ H(\lambda)u_2\rangle|}{\|u_2\|^2} \ \leqslant\ \widehat{\sigma}$ readily hold for any $u_1, u_2 \in S$ since $-\bar{\sigma}, -\widehat{\sigma}$ are the two eigenvalues of $H(\lambda)$ when restricted to the subspace S. Considering $u_1 = (l(\lambda))^2 x + \Big(\|x\|^2 - \frac{\epsilon\langle\lambda\ ,\ x\rangle}{\|\lambda\|}\Big)\lambda$ and $u_2 = \lambda - \frac{\|\lambda\|^2}{\langle\lambda\ ,\ x\rangle}$, it is easily verified that $\Big\langle x - \frac{\epsilon}{\|\lambda\|}\lambda\ ,\ u_1\Big\rangle = 0$, and $\langle\lambda\ ,\ u_2\rangle = 0$. Moreover, we also get the inequalities

$$\begin{cases} \langle u_1\ ,\ H(\lambda)u_1\rangle \ =\ \dfrac{-\epsilon}{2l(\lambda)\,\|\lambda\|}\,\|u_1\|^2 + \ \dfrac{\epsilon}{2l(\lambda)\,\|\lambda\|}\,\dfrac{|\langle\lambda\ ,\ u_1\rangle|^2}{\|\lambda\|^2} & \geqslant \dfrac{-\epsilon}{2l(\lambda)\,\|\lambda\|}\,\|u_1\|^2\,, \\ \langle u_2\ ,\ H(\lambda)u_2\rangle \ =\ \dfrac{-\epsilon}{2l(\lambda)\,\|\lambda\|}\,\|u_2\|^2 - \ \dfrac{1}{4(l(\lambda))^3}\left|\Big\langle x - \dfrac{\epsilon}{\|\lambda\|}\lambda\ ,\ u_2\Big\rangle\right|^2 & \leqslant \dfrac{-\epsilon}{2l(\lambda)\,\|\lambda\|}\,\|u_2\|^2\,. \end{cases}$$

Since the hessian $H(\lambda)$ is negative semidefinite, the inequalities (68) are obtained at once.

To complete the proof for the case when λ and x are linearly dependent, we first see that the expressions in (66) are continuous w.r.t. λ. Also, it is evident that the mapping $\Lambda \ni \lambda \longmapsto H(\lambda)$ is continuous. Since the eigenvalues of a matrix vary continuously, these two limits must be the same. The proof is now complete. ■

Proof [Lemma 2.16] We begin by first establishing that $\bar{\sigma}(H(\lambda))$ and $\widehat{\sigma}(H(\lambda))$ as given in (5.10) satisfy the inequalities

$$\alpha' \ \leqslant\ 2\bar{\sigma}(H(\lambda)) \ \leqslant\ 2\widehat{\sigma}(H(\lambda)) \ \leqslant\ \beta' \quad \text{for every } \lambda \in \Lambda(\bar{\eta}, B). \tag{69}$$

Since the mapping $\Lambda \ni \lambda \longmapsto H(\lambda)$ is concave, all the eigenvalues of the hessian $H(\lambda)$ are non-positive (more importantly, real-valued). Thus, $\frac{8\epsilon(l(\lambda))^6}{\left(\|x\|^2 - \epsilon^2\right)^2\|\lambda\|^3} \ \leqslant\ 1$ since the square

root term in (5.10) must be real valued. To prove the lower bound for $\bar{\sigma}(H(\lambda))$ in (69), we use the inequality that $\sqrt{1-\theta^2} < 1 - \frac{\theta^2}{2}$ for every $\theta \in [0,1]$. Thereby,

$$\begin{aligned}\bar{\sigma}(H(\lambda)) \;&>\; \frac{\big(\,\|x\|^2 - \epsilon^2\big)}{8(l(\lambda))^3}\left(\frac{4\epsilon(l(\lambda))^6}{\big(\,\|x\|^2-\epsilon^2\big)^2\,\|\lambda\|^3}\right)\\ &=\; \frac{\epsilon}{2\big(\,\|x\|^2-\epsilon^2\big)}\left(\frac{(l(\lambda))}{\|\lambda\|}\right)^3 \;\geqslant\; \frac{\epsilon}{2\big(\,\|x\|^2-\epsilon^2\big)}\left(\frac{\bar{\eta}}{B}\right)^3\\ &=\; \big(1/2\big)\alpha' \;\text{ for all } \lambda \in \Lambda(\bar{\eta}, B).\end{aligned}$$

For $\widehat{\sigma}(H(\lambda))$, using the inequality $\sqrt{1-\theta^2} < 1$ for $\theta \in [0,1]$, we immediately get

$$\widehat{\sigma}(H(\lambda)) \leqslant \frac{\big(\,\|x\|^2-\epsilon^2\big)}{8(l(\lambda))^3}\,2 \;\leqslant\; \frac{\big(\,\|x\|^2-\epsilon^2\big)}{4\bar{\eta}^3} \;=\; \big(1/2\big)\beta' \quad \text{for all } \lambda \in \Lambda(\bar{\eta}, B).$$

Since $\Lambda(\bar{\eta}, B) \subset \Lambda$, for every $\bar{\eta} \leqslant c^*$ and $B > 0$, it follows from assertion (i) of Lemma 5.10 that the mapping $\Lambda(\bar{\eta}, B) \ni \lambda \longmapsto -2l(\lambda)$ is also twice continuously differentiable. with the Hessian evaluated at λ being $-2H(\lambda)$. Moreover, the smallest and largest eigenvalues of this hessian are $2\bar{\sigma}(H(\lambda))$ and $2\widehat{\sigma}(H(\lambda))$ respectively. In view of the inequalities (69), we see that the minimum eigenvalue of the hessian of the map $\Lambda(\bar{\eta}, B) \ni \lambda \longmapsto -2l(\lambda)$ is bounded below by $\alpha'(\bar{\eta}, B)$ (and the maximum eigenvalue is bounded above by $\beta'(\bar{\eta})$), uniformly over $\lambda \in \Lambda(\bar{\eta}, B)$. Thus, the mapping $\Lambda(\bar{\eta}, B) \ni \lambda \longmapsto -2l(\lambda)$ is α'-strongly convex and β'-smooth. This completes the proof of the lemma. ■

5.3 Proofs for step-size selection

Proof [Proposition 3.3] For a given h, d, we first observe that the mapping $[0,1] \ni \gamma \longmapsto \eta(h+\gamma d)$ is convex since the mapping $h \longmapsto \eta\,(h)$ is convex. Consequently, the first order optimality conditions for (31) are necessary and sufficient. Now, denoting $\eta_\gamma := \eta(h+\gamma d)$, we have $\frac{\partial \eta_\gamma}{\partial \gamma} = \langle \nabla\eta(h+\gamma d)\;,\; d\rangle$, and the first order optimality conditions read

1. If $\langle \nabla\eta(h)\;,\; d\rangle = \left.\frac{\partial \eta_\gamma}{\partial \gamma}\right|_{\gamma=0} \geqslant 0$, then $\gamma^* = 0$
2. If $\nabla\eta(h+d)$ exists, and $\langle \nabla\eta(h+d)\;,\; d\rangle = \left.\frac{\partial \eta_\gamma}{\partial \gamma}\right|_{\gamma=1} \leqslant 0$, then $\gamma^* = 1$

Substituting for $\nabla\eta(h+\gamma d)$ from (9) and observing that

$$\operatorname{sgn}\langle \nabla\eta(h+\gamma d)\;,\; d\rangle = -\operatorname{sgn}\langle x - \eta_\gamma \phi(h+\gamma d)\;,\; \phi(d)\rangle\,,$$

the first order optimality conditions are equivalently written as

1. $\langle x\;,\; \phi(d)\rangle \leqslant \eta\,(h)\langle \phi(h)\;,\; \phi(d)\rangle$, implies that $\langle \nabla\eta(h)\;,\; d\rangle \geqslant 0$, and thus $\gamma^* = 0$.
2. $\langle x\;,\; \phi(d)\rangle \geqslant \eta(h+d)\langle \phi(h+d)\;,\; \phi(d)\rangle$ implies that $\langle \nabla\eta(h+d)\;,\; d\rangle \leqslant 0$, and thus $\gamma^* = 1$.

If both the above conditions fail, then we know that there exists $\gamma^* \in (0,1)$ such that $\left.\frac{\partial \eta_\gamma}{\partial \gamma}\right|_{\gamma=0} = 0$. Equivalently, we have $0 = \langle \nabla\eta(h+\gamma^* d)\;,\; d\rangle = \big\langle x - \eta_{\gamma^*}\phi(h+\gamma^* d)\;,\; \phi(d)\big\rangle$.

Substituting for $\eta_{\gamma*}$ from (8) and simplifying gives the following equation in γ^*

$$\frac{\langle \phi(h+\gamma^* d)\ ,\ \phi(d)\rangle}{\langle x\ ,\ \phi(d)\rangle} = \frac{1}{\eta_{\gamma*}} = \frac{\langle x\ ,\ \phi(h+\gamma^* d)\rangle + \|\phi(h+\gamma^* d)\|\sqrt{e(h+\gamma^* d)-\epsilon^2}}{\left(\|x\|^2-\epsilon^2\right)}.$$

Rearranging terms, and substituting for $e(h+\gamma^* d)$ from (6), results in the folllowing equation

$$\begin{aligned}\frac{\langle \phi(h+\gamma^* d)\ ,\ \phi(d)\rangle}{\langle x\ ,\ \phi(d)\rangle} - \frac{\langle x\ ,\ \phi(h+\gamma^* d)\rangle}{\left(\|x\|^2-\epsilon^2\right)} \\ = \frac{\sqrt{\langle x\ ,\ \phi(h+\gamma^* d)\rangle^2 - \|\phi(h+\gamma^* d)\|^2\left(\|x\|^2-\epsilon^2\right)}}{\left(\|x\|^2-\epsilon^2\right)}.\end{aligned} \tag{70}$$

Finally, on squaring both sides of (70), we obtain the equation $a{\gamma^*}^2 + 2b\gamma^* + c = 0$, for values of a, b, c given in (32).

If $a = 0$, $\gamma^* = {}^{-c}\!/_{2b}$ is the only solution. Whereas, if $a \neq 0$, it must be observed that out of the two roots of the quadratic equation $a{\gamma^*}^2 + 2b\gamma^* + c = 0$, one satisfies (70) and the other satisfies

$$\begin{aligned}\frac{\langle \phi(h+\gamma d)\ ,\ \phi(d)\rangle}{\langle x\ ,\ \phi(d)\rangle} - \frac{\langle x\ ,\ \phi(h+\gamma d)\rangle}{\left(\|x\|^2-\epsilon^2\right)} \\ = -\frac{\sqrt{\langle x\ ,\ \phi(h+\gamma d)\rangle^2 - \|\phi(h+\gamma d)\|^2\left(\|x\|^2-\epsilon^2\right)}}{\left(\|x\|^2-\epsilon^2\right)}.\end{aligned}$$

Thus, the correct root of the quadratic equation can be picked by ensuring the criterion

$$0 \leqslant \frac{\langle \phi(h+\gamma d)\ ,\ \phi(d)\rangle}{\langle x\ ,\ \phi(d)\rangle} - \frac{\langle x\ ,\ \phi(h+\gamma d)\rangle}{\left(\|x\|^2-\epsilon^2\right)}.$$

The proof of the proposition is now complete. ■

References

M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo. An Augmented Lagrangian Approach to the Constrained Optimization Formulation of Imaging Inverse Problems. *IEEE Transactions On Image Processing*, 2009. doi: 10.1109/TIP.2010.2076294.

M. Aharon, M. Elad, and A. Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11): 4311–4322, 2006. ISSN 1053587X. doi: 10.1109/TSP.2006.881199.

A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009. ISSN 19364954. doi: 10.1137/080716542.

D. P Bertsekas. *Control of uncertain systems with a set-membership description of the uncertainty.* PhD thesis, Massachusetts Institute of Technology, 1971.

S. Boyd, L. Xiao, and A. Mutapcic. Subgradient methods. *Notesfor EE392o*, 2003.

E. J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

E. J. Candès and M. B. Wakin. An introduction to compressive sampling [a sensing/sampling paradigm that goes against the common knowledge in data acquisition]. *IEEE signal processing magazine*, 25(2):21–30, 2008.

E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9:717–772, 12 2009. ISSN 16153375. doi: 10.1007/s10208-009-9045-5.

A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. Technical report, HAL open science, 2010.

A. Chambolle and T. Pock. On the ergodic convergence rates of a first-order primal–dual algorithm. *Mathematical Programming*, 159(1-2):253–287, 9 2016. ISSN 14364646. doi: 10.1007/s10107-015-0957-3.

V. Chandrasekaran, B. Recht, P.A. Parrilo, and A. S. Willsky. The convex geometry of linear inverse problems. *Foundations of Computational mathematics*, 12(6):805–849, 2012.

D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52:1289–1306, 4 2006a. ISSN 00189448. doi: 10.1109/TIT.2006.871582.

D. L Donoho. For Most Large Underdetermined Systems of Linear Equations the Minimal 1-norm Solution Is Also the Sparsest Solution. *Communications on Pure and Applied Mathematics*, 59:797–829, 2006b. doi: 10.1002/cpa.20132.

J. Duchi, Shai Shalev-Schwartz, Yoram Singer, and Tushar CHandra. Efficient Projections onto the l1-Ball for Learning in High Dimensions. Technical report, Proceedings of the 25th International Conference on Machine Learning, 2008.

M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image processing*, 15(12):3736–3745, 2006.

M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval Research Logistics*, 3 1956.

S. Gleichman and Y. C. Eldar. Blind Compressed Sensing. *IEEE Transactions on Information Theory*, 57(10):6958–6975, 10 2011. doi: 10.1109/TIT.2011.2165821.

M. Jaggi. Revisiting Frank-Wolfe: Projection-Free Sparse Convex Optimization. Technical report, Ecole Polytechnique, 2013.

M. Nagahara, D. E. Quevedo, and D. Nešić. Maximum hands-off control: a paradigm of control effort minimization. *IEEE Transactions on Automatic Control*, 61(3):735–747, 2015.

B. A. Olshausen and D. J. Fieldt. Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1 ? Technical Report 23, 1997.

B. Recht, M. Fazel, and P.A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.

M. R. Sheriff and Debasish Chatterjee. Novel min-max reformulations of Linear Inverse Problems. *arXiv preprint arXiv:2007.02448.*, 7 2020.

M.R. Sheriff, F.F. Redel, and P. Mohajerin Esfahani. The Fast Linear Inverse Problem Solver (FLIPS). `https://github.com/MRayyanS/FLIPS`, 2022.

M. Yaghoobi and M. E. Davies. Compressible dictionary learning for fast sparse approximations. In *IEEE Workshop on Statistical Signal Processing Proceedings*, pages 662–665, 2009. ISBN 9781424427109. doi: 10.1109/SSP.2009.5278490.

W. H. Yang. On generalized holder inequality. *Nonlinear Analysis Theory, Methods and Applications*, 16:489–498, 1991.

Ye. E. Nesterov. A method of solving a convex programming problem with convergence rate O(1/kˆ2). *Soviet Math dokl.*, 27(2), 1983.