Geometry of First-Order Methods and Adaptive Acceleration

Clarice Poon[*] Jingwei Liang[†]

Abstract

First-order operator splitting methods are ubiquitous among many fields through science and engineering, such as inverse problems, signal/image processing, statistics, data science and machine learning, to name a few. In this paper, we study a geometry property of first-order methods when applying to solve non-smooth optimization problems. With the tool "partial smoothness", we design a framework to analyze the trajectory of the fixed-point sequence generated by first-order methods and show that locally the fixed-point sequence settles onto a regular trajectory such as straight line or spiral. Based on this finding, we discuss the limitation of current widely used "inertial acceleration" technique, and propose a trajectory following adaptive acceleration algorithm. Global convergence is established for the proposed acceleration scheme based on the perturbation of fixed-point iteration. Locally, we first build connections between the acceleration scheme and the well studied "vector extrapolation technique" in the field of numerical analysis, and then discuss acceleration guarantees of the proposed acceleration scheme. Moreover, our result provides a geometric interpretation of vector extrapolation. Numeric experiments on various first-order methods are provided to demonstrate the advantage of the proposed adaptive acceleration scheme.

Key words. Non-smooth optimization, first-order methods, inertial acceleration, partial smoothness, finite activity identification, trajectory of sequence, vector extrapolation.

AMS subject classifications. 65B05, 65K05, 65K10, 90C25, 90C31.

1 Introduction

Non-smooth optimization is an active research area of modern optimization, which aims to find solutions of structured problems that are the sum of smooth and non-smooth functions, possibly under constraints and composition with (non)linear operators. It plays a fundamental role in various fields through science and engineering, such as inverse problems, signal/image processing, compressed sensing, statistics, data science and machine learning, etc. In the literature, numerical schemes, typically first-order (operator/proximal splitting) methods, have been designed to solve non-smooth optimization problems. Over the past decades, driven by the real-world problems arising from the aforementioned fields, non-smooth optimization and first-order methods have experienced tremendous growth and success, especially in large-scale problems. However, despite the huge success, first-order methods suffer a significant drawback: slow speed of convergence, which has made them the bottleneck of solving today's even larger-scale problems. With the increasing size of data sets and growing complexity of mathematical models of real-world problems, the need for novel fast and low computational cost algorithms is becoming increasingly strong.

In this paper, we denote **FoM** the class of first-order methods and $\mathscr{F} \in \textbf{FoM}$ a first-order method, for instance the proximal gradient descent; See Example 1.1. The iteration of $\mathscr{F}$ usually can be (re)formulated as a fixed-point iteration in a proper real Hilbert space $\mathscr{H}$ of the form

$$z_{k+1} = \mathscr{F}(z_k), \tag{1.1}$$

where $\{z_k\}_{k\in\mathbb{N}}$ is the fixed-point iterates that converges to $z^\star \in \mathrm{fix}(\mathscr{F})$ with $\mathrm{fix}(\mathscr{F}) \overset{\text{def}}{=} \{z \in \mathscr{H} : z = \mathscr{F}(z)\}$ which is supposed to be non-empty. We refer to [7] for more detailed accountant of fixed-point theory.

[*] Department of Mathematics, University of Bath, Bath UK. E-mail: cmhsp20@bath.ac.uk.

[†] DAMTP, University of Cambridge, Cambridge UK. E-mail: jl993@cam.ac.uk.

1.1 Acceleration of first-order methods

In the literature, numerous approaches are proposed to accelerate first-order methods, among them, the "inertial technique" and over-relaxation are probably the most widely used two. Both these approaches belong to the realm of *extrapolation technique*. Let $\mathscr{F}$ be the first-order method in (1.1), a general combination of extrapolation and $\mathscr{F}$ takes the following form

$$\begin{aligned} \bar{z}_k &= \mathscr{E}(\bar{z}_{k-1}, z_k, z_{k-1}, \ldots), \\ z_{k+1} &= \mathscr{F}(\bar{z}_k, z_k, z_{k-1}, \ldots), \end{aligned} \tag{1.2}$$

where $\mathscr{E}$ is the extrapolation step that computes the point $\bar{z}_k$ based on its previous point $\bar{z}_{k-1}$ and the history of z_k including $\{z_k, z_{k-1}, ..\}$. In what follows we present a brief overview of inertial technique and over-relaxation.

1.1.1 Inertial acceleration

The very first inertial scheme is the "heavy-ball method" [64] proposed by Polyak which can significantly speed-up the performance of gradient descent, particularly when the problem is strongly convex and twice differentiable. The theoretical foundation of inertial acceleration was due to Nesterov, in [59] he showed that a different combination of inertial and gradient descent can improve the $O(1/k)$ convergence rate of objective function value to $O(1/k^2)$. This result was further extended to the non-smooth case by Beck and Teboulle in [11] where they proposed the "fast iterative shrinkage-thresholding algorithm", a.k.a FISTA for speeding up Forward–Backward splitting method [51] (*i.e.* the proximal gradient descent). Note that, gradient descent and its proximal version are descent methods, that is the objective function value along the iteration is monotonically non-increasing along iteration[1].

Over the years, the huge success of the accelerated (proximal) gradient descent schemes has motivated people to extend the inertial acceleration to other first-order methods. Let $\mathscr{F}$ be the first-order method in (1.1), a generic inertial version of $\mathscr{F}$ would read

$$\begin{aligned} \bar{z}_k &= z_k + a_k(z_k - z_{k-1}), \\ z_{k+1} &= \mathscr{F}(\bar{z}_k, z_k). \end{aligned} \tag{1.3}$$

The inertial scheme first extrapolate a point $\bar{z}_k$ along the direction of $z_k - z_{k-1}$, and then update the next z_{k+1} based on $\bar{z}_k$ with or without z_k. The formulation (1.3) abstracts many existing inertial schemes in the literature, below is an example of gradient descent.

Example 1.1 (Gradient descent). Consider an unconstrained smooth minimization problem, $\min_{x\in\mathbb{R}^n} F(x)$ where $F : \mathbb{R}^n \to \mathbb{R}$ is proper convex differentiable with gradient ∇F being L-Lipschitz continuous. The iteration of gradient descent reads (for this case we use x_k instead of z_k)

$$x_{k+1} = x_k - \gamma\nabla F(x_k),$$

where $\gamma \in]0, 2/L[$ is the step-size. The fixed-point operator of gradient descent reads $\mathscr{F} \overset{\text{def}}{=} \mathrm{Id} - \gamma\nabla F$. The "heavy-ball method" [64] takes the following form of iteration

$$\begin{aligned} \bar{x}_k &= x_k + a_k(x_k - x_{k-1}), \\ x_{k+1} &= \bar{x}_k - \gamma\nabla F(x_k), \end{aligned} \tag{1.4}$$

where $a_k \in [0, 1]$ is the inertial parameter. If we further replace $\nabla F(x_k)$ with $\nabla F(\bar{x}_k)$, and compute a_k via $t_k = \frac{1+\sqrt{1+4t_{k-1}^2}}{2}$, $a_k = \frac{t_{k-1}-1}{t_k}$ with $t_0 = 1$, then (1.4) becomes the scheme of [59] which achieves $O(1/k^2)$ convergence rate for $F(x_k) - F(x^\star)$ where $x^\star$ is a global minimizer of F.

Other examples of inertial first-order methods include: the inertial versions of Proximal Point Algorithm

[1] Descent methods include gradient descent and its proximal version (a.k.a. Forward–Backward splitting) and proximal point algorithm, note that the problem does not necessarily to be convex [6] for the method to be descent. Other first-order methods, such as Douglas–Rachford/ADMM and Primal–Dual splitting methods, are non-descent in general.

[4, 3], Forward–Backward splitting [58, 52, 45], Douglas–Rachford splitting [16] and inertial Primal–Dual splitting [14], or in general the inertial version of Krasnosel'skiĭ-Mann fixed-point iteration [53, 32] which covers many of the inertial first-order methods as special cases. However, despite its overwhelming popularity, the combination of inertial technique and first-order methods suffers several drawbacks

- **Restricted parameter choices** Unlike the elegant inertial (proximal) gradient descent methods [64, 59, 11], the choices of (inertial) parameters for general inertial first-order methods, *e.g.* [16, 52, 53, 32], are quite restricted and complicated.
- **Complicated convergence proof** For inertial (proximal) gradient descent methods [64, 59, 11], a Lyapunov stability function can be found easily, even in the non-convex case, as (proximal) gradient descent is a descent method. Things become much more complicated for other first-order methods as they are non-decent, and Lyapunov functions can only be obtained under stronger assumptions or does not exist at all. As a result, the convergence proof becomes more complicated, which is also another reason of restricted parameter choices.
- **Lack of acceleration guarantees** As these methods are not descent, there are limited acceleration guarantees, unless stronger assumptions, such as smoothness or strong convexity, are imposed. Examples of inertial schemes failing to provide acceleration can be easily found when no stronger assumptions are available; See Section 4 for examples, and also [66] and [45, Chapter 4.5].

Finally, it is worth mentioning that, in the literature, most inertial schemes consider only the momentum created two past points, namely $z_k - z_{k-1}$. For certain cases, use the momentum of more than two points could be beneficial, see Section 4 for example. This is mentioned in [64], and related work can be found in [45, 31].

1.1.2 Over-relaxation

In the field of fixed-point theory, another popular approach to accelerate convergence is the *over-relaxation* which is the generalization of the successive over relaxation for linear systems. For the fixed-point iteration (1.1), the relaxation of it reads

$$z_{k+1} = z_k + \lambda_k\big(\mathscr{F}(z_k) - z_k\big), \tag{1.5}$$

where $\lambda_k \in]0, \bar{\lambda}]$ is the relaxation parameter and $\bar{\lambda}$ is the upper bound of λ determined by the property of $\mathscr{F}$. For example $\bar{\lambda} = \frac{1}{\alpha}$ when $\mathscr{F}$ is so-called α-averaged non-expansive; see Definition 2.2 and [7] for more detailed discussions.

When $\bar{\lambda} > 1$ and $\lambda \in]1, \bar{\lambda}]$, (1.5) is the over-relaxed version of $\mathscr{F}$. Below we briefly show that over-relaxed $\mathscr{F}$ is equivalent to an inertial version of (1.1) which is a special case of (1.3). Denote $a_k = \lambda_k - 1$, then we can rewrite (1.5) as[2]

$$\begin{aligned} \bar{z}_k &= z_k + a_k(z_k - \bar{z}_{k-1}), \\ z_{k+1} &= \mathscr{F}(\bar{z}_k). \end{aligned} \tag{1.6}$$

In stead of extrapolating a point along the direction $z_k - z_{k-1}$, relaxation uses $z_k - \bar{z}_{k-1}$. Over-relaxation can also significantly improved the convergence speed of (1.1), such as the over-relaxed projection based algorithms for feasibility problems [38, 9]. However, same as the inertial scheme (1.3), over-relaxation is not guarantee to provide acceleration. For instance, in [9, 48], the authors showed that the optimal λ_k for Douglas–Rachford splitting when applied to (locally) polyhedral problem is 1, that is no relaxation provides the best performance.

Generally speaking, over-relaxation suffers the same problem as inertial schemes, that its acceleration guarantees are methods and problems dependent:

[2] Strictly speaking, (1.5) is equivalent to

$$\begin{aligned} z_k &= \bar{z}_k + a_k(\bar{z}_k - z_{k-1}), \\ \bar{z}_{k+1} &= \mathscr{F}(z_k). \end{aligned}$$

We switch z_k and $\bar{z}_k$ in order to comply with (1.2).

- For descent methods, *e.g.* Forward–Backward splitting, owing to the result of [47], it can be shown that locally over-relaxation can provide acceleration.
- While for other algorithms, such as Douglas–Rachford splitting, the performance of over-relaxation depends on the problem to solve and parameters of the algorithm [9, 48].

1.2 Main contributions

In this paper, motivated by the behavior of inertial first-order methods and over-relaxation scheme, we present a systematic study on the geometric properties of first-order methods and their acceleration. We first show that the performance of inertial and relaxation is determined by the trajectory of the generated fixed-point sequence $\{z_k\}_{k\in\mathbb{N}}$, that different trajectories result in different outcomes (see Section 3). When considering non-smooth optimization, we present a unified framework for analyzing the local trajectory of the fixed-point sequence of first-order methods. Based on this finding, we propose a generic trajectory following linear prediction scheme for accelerating first-order methods. More precisely, our contributions in this paper are summarized below.

Geometry of FoM via trajectory of fixed-point sequences In the literature of first-order methods, numerous first-order operator splitting methods are proposed which is the consequence of the structures of the optimization problems to solve. However, the study on the structure of first-order methods is rather limited, which is mainly due to the non-linearity of the iteration. In this paper, by focusing on non-smooth optimization, with the help of "partial smoothness" (Definition 2.6), in Section 3 we propose a generic framework for analyzing the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ generated by the fixed-point iteration (1.1) (Section 3.1). More precisely, we show that $\mathscr{F}$ can be linearized locally around the solution along some C^2-smooth manifold(s), up to residuals. This means there exists a square matrix $M_{\mathscr{F}}$ such that

$$z_{k+1} - z_k = M_{\mathscr{F}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|).$$

Based on the spectral properties of $M_{\mathscr{F}}$, we show that different first-order methods admit different types of trajectories for the fixed-point sequence $\{z_k\}_{k\in\mathbb{N}}$:

- For (proximal) gradient descent, we show that the spectrum of $M_{\mathscr{F}}$ is real, as a result the eventual trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a straight line; See Section 3.2.
- For other popular first-order methods, such as Douglas–Rachford splitting and alternating direction method of multipliers (ADMM), based on the properties of the functions and parameters, we show that the leading eigenvalue of $M_{\mathscr{F}}$ can be either real or complex, and the eventual trajectory of $\{z_k\}_{k\in\mathbb{N}}$ could be either a straight line (real leading eigenvalue) or a spiral (complex leading eigenvalue); See Section 3.3. For Primal–Dual splitting methods, the leading eigenvalue of M is complex and the eventual trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a spiral; see Section 3.4.

Limitation of inertial and over-relaxation The trajectory of first-order methods allows us to analyze the limitations of inertial technique and over-relaxation. In Section 4, based on examples of Douglas–Rachford splitting method, we show that for inertial

- When the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a *straight line*, then inertial can provide substantial acceleration.
- When the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a *logarithmic spiral*, we show that inertial will always fail to provide acceleration, and one should not consider relaxation neither.
- When the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a *elliptical spiral*, inertial and over-relaxation can provide acceleration under proper implementation.

An adaptive acceleration via linear prediction The limitation of inertial and over-relaxation techniques, particularly their failures, implies that the correct acceleration scheme should be able to adapt to the trajectory of the underlying sequence, which is another core contribution of this work. By exploiting the eventual regularity, *i.e.* either straight line or spiral, of the trajectory of $\{z_k\}_{k\in\mathbb{N}}$, in Section 5 we propose an adaptive linear prediction scheme for accelerating first-order methods which is able to following the trajectory of the fixed-point sequence. Global convergence based on perturbation of fixed-point iteration is provided. Local acceleration guarantees are also provided for the proposed scheme, based on the connections with existing vector extrapolation techniques.

Our proposed linear prediction scheme belongs to the realm of vector extrapolation method, while our derivation provides an alternative geometric interpretation for polynomial extrapolation methods such as minimal polynomial extrapolation (MPE) [21] and reduced rank extrapolation (RRE) [34, 55]. Our linear prediction bridges the gap between inertial schemes and polynomial extrapolation methods. Moreover, our geometric interpretation of linear prediction provides insights on how to enhance the robustness and performance of extrapolation methods.

1.3 Related work

Over the past decades, owing to the tremendous success of inertial acceleration [59, 11], the inertial technique has been widely adapted to accelerate other first-order algorithms. For example inertial Douglas–Rachford and alternating direction method of multipliers (ADMM) [15, 62, 41, 36], inertial Primal–Dual splitting [25, 45]. In terms of Krasnosel'skiĭ-Mann fixed-point iteration, the inertial of it are also studied in the literature such as [53, 30]. Multi-step inertial schemes, *i.e.* using the momentum created by more than two past points, are also considered in the literature, see for instance [31, 45]. However, for most of these works, to ensure acceleration guarantees of inertial, stronger assumptions are needed, such as Lipschitz continuity or strong convexity, see [36] for ADMM. When it comes to general non-smooth problems, some of them would fail to provide acceleration. Moreover, as discussed in [45, Chapter 4], for certain problems and algorithms, such as basis pursuit and Douglas–Rachford splitting algorithm, only multi-step inertial scheme can provide acceleration.

For more generic acceleration techniques, there are extensive works in numerical analysis on the topic of convergence acceleration for sequences. Given an arbitrary sequence $\{z_k\}_{k\in\mathbb{N}} \subset \mathbb{R}^n$ with limit $z^\star$, the goal of convergence acceleration is to find a transformation $\mathscr{E}_k : \{z_{k-j}\}_{j=1}^q \to \bar{z}_k \in \mathbb{R}^n$ such that $\bar{z}_k$ converges faster to $z^\star$. In general, the process by which $\{z_k\}$ is generated is unknown, q is chosen to be a small integer, and $\bar{z}_k$ is referred to as the extrapolation of z_k. Some of the best known examples include Richardson's extrapolation [68], the Δ^2-process of Aitken [2] and Shank's algorithm [72]. We refer to the article [17] and books [19, 73] for a detailed historical perspective on the development of these techniques. Much of the works on the extrapolation of vector sequences was initiated by Wynn [79] who generalized the work of Shank to vector sequences. In Section 6, the formulation of some of these methods are provided. In particular, minimal polynomial extrapolation (MPE) [21] and Reduced Rank Extrapolation (RRE) [34, 55] (which is also a variant of Anderson acceleration developed independently in [5]), which are particularly relevant to this present work.

More recently, there has been a series of work on a regularized version of RRE [71, 70, 12]. As mentioned in [74], the stability of vector extrapolation techniques depend on the stability of computing the extrapolation coefficients, to address this instability, [74] proposed to apply Tikhonov regularization when computing the extrapolation coefficients. We remark however the regularization parameter in these works rely on a grid search based on objective function which is only doable for descent methods. The contributions of this work is therefore different: First, we are interested in more general optimization problems (such as finding the intersection of convex sets or basis pursuit) where there may not be an objective to minimize. Second, we directly handle the non-smoothness of optimization problems by studying the eventual trajectories of the generated sequences.

Paper organization The organization of the paper is as following: necessary notations and definitions are collected in Section 2. In Section 3, we propose a generic framework for analyzing the trajectory of first-order methods, Forward–Backward, Douglas–Rachford and Primal–Dual splitting methods are discussed in details. The limitations of inertial technique and over-relaxation when applied to non-descent methods are discussed in Section 4. The trajectory motivated linear prediction acceleration scheme is described in Section 5, where global convergence is also provided. In Section 6, by connecting linear prediction with existing polynomial extrapolation methods, local acceleration guarantees are provided. Numerical experiments are presented in Section 7. Trajectory of linear system and proofs of main propositions are collected in the appendix.

2 Mathematical background

Throughout the paper, $\mathbb{R}^n$ is a n-dimensional Euclidean space equipped with scalar inner product $\langle\cdot,\cdot\rangle$ and associated norm $\|\cdot\|$. Id denotes the identity operator on $\mathbb{R}^n$. $\Gamma_0(\mathbb{R}^n)$ denotes the class of proper convex and lower semi-continuous functions on $\mathbb{R}^n$.

2.1 Convex and set-valued analysis

For a nonempty convex set $C \subset \mathbb{R}^n$, denote by $\mathrm{aff}(C)$ its affine hull and by $\mathrm{par}(C)$ the smallest subspace parallel to $\mathrm{aff}(C)$. Denote ι_C the indicator function of C, $\mathcal{N}_C$ the associated normal cone operator and $\mathcal{P}_C$ the orthogonal projection operator on C.

The sub-differential of a proper convex and lower semi-continuous function $R \in \Gamma_0(\mathbb{R}^n)$ is the set-valued operator defined by $\partial R : \mathbb{R}^n \rightrightarrows \mathbb{R}^n, x \mapsto \big\{g \in \mathbb{R}^n | R(x') \geq R(x) + \langle g, x'-x\rangle, \forall x' \in \mathbb{R}^n\big\}$. Let $\gamma > 0$, the proximity operator or proximal mapping, of R is defined by

$$\mathrm{prox}_{\gamma R}(\cdot) \overset{\mathrm{def}}{=} \mathrm{argmin}_{x\in\mathbb{R}^n}\, \gamma R(x) + \tfrac{1}{2}\|x-\cdot\|^2.$$

The Fenchel conjugate, or simply conjugate, of R is defined by $R^*(v) \overset{\mathrm{def}}{=} \sup_{x\in\mathbb{R}^n}\left(\langle x,\, v\rangle - R(x)\right)$.

Definition 2.1 (Monotone operator). A set-valued mapping $A : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be monotone if, given any $x_1, x_2 \in \mathbb{R}^n$ there holds

$$\langle x_1 - x_2,\, v_1 - v_2\rangle \geq 0,\ \forall v_1 \in A(x_1) \text{ and } v_2 \in A(x_2).$$

It is maximal monotone if its $\mathrm{gph}\,(A) \overset{\mathrm{def}}{=} \{(x,v) \in \mathbb{R}^n \times \mathbb{R}^n | v \in A(x)\}$ can not be contained in the graph of any other monotone operators.

For a maximal monotone operator A, $(\mathrm{Id}+A)^{-1}$ denotes its resolvent. It is known that for function $R \in \Gamma_0(\mathbb{R}^n)$, its sub-differential ∂R is maximal monotone [69], and that $\mathrm{prox}_R = (\mathrm{Id}+\partial R)^{-1}$.

Definition 2.2 (Non-expansive operator). An operator $\mathscr{F} : \mathbb{R}^n \to \mathbb{R}^n$ is non-expansive if

$$\|\mathscr{F}(x) - \mathscr{F}(y)\| \leq \|x-y\|,\ \ \forall x,y \in \mathbb{R}^n.$$

That is, $\mathscr{F}$ is 1-Lipschitz continuous. For any $\alpha \in]0,1[$, $\mathscr{F}$ is called α-averaged if there exists a non-expansive operator $\mathscr{F}'$ such that $\mathscr{F} = \alpha\mathscr{F}' + (1-\alpha)\mathrm{Id}$.

The fixed points of non-expansive operators in general are not available explicitly. To find them, one has to apply certain iterative procedures, one of the most-known is the Krasnosel'skiĭ-Mann iteration [42, 54].

Definition 2.3 (Krasnosel'skiĭ-Mann iteration). Let $\mathscr{F} : \mathbb{R}^n \to \mathbb{R}^n$ be a non-expansive operator such that $\mathrm{fix}(\mathscr{F}) \neq \emptyset$. Let $\lambda_k \in [0,1]$ and choose $x_0 \in \mathbb{R}^n$ arbitrarily, the Krasnosel'skiĭ-Mann iteration of $\mathscr{F}$ reads

$$z_{k+1} = z_k + \lambda_k(\mathscr{F}(z_k) - z_k).$$

Moreover, if $\lambda_k \in [0,1]$ is such that $\sum_{k\in\mathbb{N}} \lambda_k(1-\lambda_k) = +\infty$, then $\{z_k\}_{k\in\mathbb{N}}$ converges to a point in $\mathrm{fix}(\mathscr{F})$ [7].

When $\mathscr{F}$ is α-averaged, the upper bound of λ_k becomes $\frac{1}{\alpha}$, and the condition needed for convergence of Krasnosel'skiĭ-Mann iteration changes to $\sum_{k\in\mathbb{N}} \lambda_k(\frac{1}{\alpha} - \lambda_k) = +\infty$.

2.2 Angle between subspaces

Let T_1, T_2 be two subspaces, and without the loss of generality, assume $1 \leq p \overset{\mathrm{def}}{=} \dim(T_1) \leq q \overset{\mathrm{def}}{=} \dim(T_2) \leq n-1$.

Definition 2.4 (Principal angles). The principal angles $\theta_k \in [0, \frac{\pi}{2}]$, $k = 1, \ldots, p$ between subspaces T_1 and T_2 are defined by, with $u_0 = v_0 \overset{\text{def}}{=} 0$, and

$$\cos(\theta_k) \overset{\text{def}}{=} \langle u_k, v_k \rangle = \max \langle u, v \rangle \text{ s.t. } u \in T_1, v \in T_2, \|u\| = 1, \|v\| = 1, \ \langle u, u_i \rangle = \langle v, v_i \rangle = 0, i = 0, \cdots, k-1.$$

The principal angles θ_k are unique and satisfy $0 \le \theta_1 \le \theta_2 \le \cdots \le \theta_p \le \pi/2$.

Definition 2.5 (Friedrichs angle). The Friedrichs angle $\theta_F \in]0, \frac{\pi}{2}]$ between T_1 and T_2 is

$$\cos(\theta_F) \overset{\text{def}}{=} \max \langle u, v \rangle \text{ s.t. } u \in T_1 \cap (T_1 \cap T_2)^{\perp}, \|u\| = 1, v \in T_2 \cap (T_1 \cap T_2)^{\perp}, \|v\| = 1.$$

Lemma 2.1 ([8]). *The Friedrichs angle is exactly θ_{d+1} where $d \overset{\text{def}}{=} \dim(T_1 \cap T_2)$. Moreover, $\theta_F > 0$.*

2.3 Partial smoothness

Let $\mathcal{M}$ be a C^2-smooth Riemannian manifold, denote $\mathcal{T}_{\mathcal{M}}(x)$ the tangent space to $\mathcal{M}$ at any point x in $\mathcal{M}$. The definition below of partial smoothness is adapted from [44] to the case of $\Gamma_0(\mathbb{R}^n)$ functions.

Definition 2.6 (Partly smooth function [44]). A function $R \in \Gamma_0(\mathbb{R}^n)$ is partly smooth at $\bar{x}$ relative to a set $\mathcal{M}_{\bar{x}}$ if $\mathcal{M}_{\bar{x}}$ is a C^2 manifold around $\bar{x}$, and:

Smoothness R restricted to $\mathcal{M}_{\bar{x}}$ is C^2-smooth around $\bar{x}$.

Sharpness The tangent space $\mathcal{T}_{\mathcal{M}_{\bar{x}}}(\bar{x}) = \mathrm{par}(\partial R(\bar{x}))^{\perp}$.

Continuity The set-valued mapping ∂R is continuous at x relative to $\mathcal{M}_{\bar{x}}$.

Loosely speaking, a partly smooth function behaves *smoothly* along the smooth manifold $\mathcal{M}_{\bar{x}}$, and *sharply* transversal to $\mathcal{M}_{\bar{x}}$. The class of partly smooth functions at $\bar{x}$ relative to $\mathcal{M}_{\bar{x}}$ is denoted as $\mathrm{PSF}_{\bar{x}}(\mathcal{M}_{\bar{x}})$. We reference [45, Chapter 5] and the references therein for popular examples of partly smooth functions which include: indicator function of partly smooth set, ℓ_1-norm, $\ell_{1,2}$-norm, ℓ_∞-norm, total variation and nuclear norm, etc. In the past few year, partial smoothness has proven to be a powerful tool for analyzing the local convergence behaviors of first-order methods [46, 45, 47, 57] when applied to non-smooth optimization.

2.4 Trajectory of sequence

Let $\{z_k\}_{k\in\mathbb{N}}$ be a train of sequence in $\mathbb{R}^n$ whose limiting point exists, by connecting all the points with line segments we obtain the trajectory of the sequence. Given k, define $v_k \overset{\text{def}}{=} z_k - z_{k-1}$ the displacement vector. To characterize the trajectory of $\{z_k\}_{k\in\mathbb{N}}$, we use the angle θ_k between two consecutive v_k, v_{k-1} which is define by

$$\theta_k \overset{\text{def}}{=} \angle(v_k, v_{k-1}) = \arccos\left(\frac{\langle v_k, v_{k-1} \rangle}{\|v_k\|\|v_{k-1}\|}\right). \tag{2.1}$$

In this paper, we are interested in three different types of trajectories, which are summarized in Table 1 below: straight line and two types of spiral (logarithmic and elliptical).

For these three types of trajectories, we have

(I) For Type I trajectory, θ_k converges to 0 which means eventually $\{z_k\}_{k\in\mathbb{N}}$ lies in a *straight line*.

(II) For Type II trajectory, instead of converging to 0, θ_k converge to some $\theta_F \in]0, \pi/2[$ implying that the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a *logarithmic spiral*. See the top view of the Type II trajectory above.

(III) For Type III trajectory, different from the former two cases, θ_k eventually oscillate in an interval, which results in an *elliptical spiral*. See the top view of the Type III trajectory above.

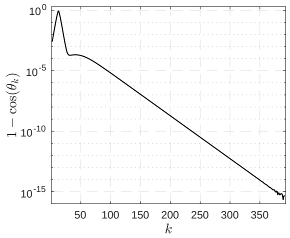
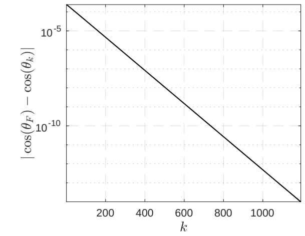
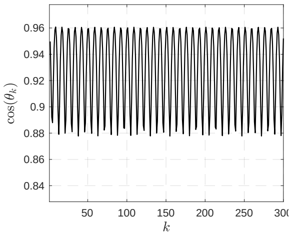
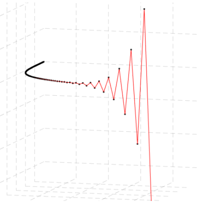
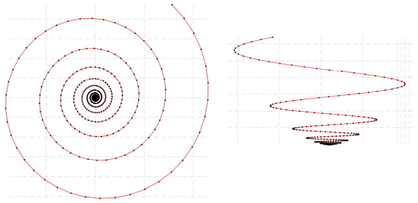
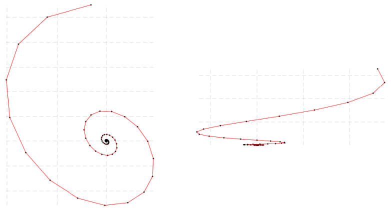
Detailed discussion on these trajectories are presented in Section A of the appendix.

Remark 2.2 (What determines the trajectory of $\{z_k\}_{k\in\mathbb{N}}$). Suppose the sequence $\{z_k\}_{k\in\mathbb{N}}$ above is generated by a liner system of the form $z_{k+1} = Mz_k$ where M is a linear matrix, and M is such that its spectral radius is strictly smaller than 1[3]. The type of trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is determined by the *leading eigenvalue* of M — real

[3] If the spectral radius of M is equal to 1, then as long as the power of M converges, *i.e.* there exists a matrix $\widetilde{M}$ such that $\widetilde{M} = \lim_{k\to+\infty} M^k$, then we can consider the leading eigenvalue of $M - \widetilde{M}$ instead of M.

leading eigenvalue leads to straight line trajectory, and complex eigenvalue leads to spiral trajectory. For the type of spiral trajectory, it relies on the further properties of the leading eigenvalue; Section A of the appendix.

Table 1: Three types of trajectory of sequence.

Type I: straight line	Type II: logarithmic spiral	Type III: elliptical spiral
$\theta_k \to 0$	$\theta_k \to \theta_F \in]0, \pi/2[$	$\theta_k \to [\underline{\theta}, \overline{\theta}] \subset]0, \pi/2[$
		
		

3 Local trajectory of first-order methods

As mentioned above, the trajectory of sequence is determined by the leading eigenvalue of the linear system, however the fixed-point operators of first-order methods in general are non-linear, which makes the study of the trajectory of the generated sequence impossible. However, when dealing with *non-smooth optimization*, locally around the solution the fixed-point operators can be linearized with respect to some smooth manifolds under the help of "partial smoothness". In this section, we present an abstract framework for analyzing the local trajectory of first-order methods and apply it to analyze several popular first-order methods. All the proofs for propositions in this section are provided in Section B of the appendix.

3.1 A framework based on partial smoothness

Recall the fixed-point iteration of first-order methods (1.1): $z_{k+1} = \mathscr{F}(z_k)$. Define the difference vector $v_k \overset{\text{def}}{=} z_k - z_{k-1}$ and the angle $\theta_k \overset{\text{def}}{=} \angle(v_k, v_{k-1})$ between v_k, v_{k-1} as in (2.1). We propose the following framework for analyzing the trajectory of sequence $\{z_k\}_{k\in\mathbb{N}}$.

A framework for analyzing local trajectory of first-order methods

1. **Convergent sequence** The iteration is convergent and $z_k \to z^\star \in \text{fix}(\mathscr{F})$.
2. **Manifold identification** Under a proper non-degenerate condition, see *e.g.* (ND_{FB}) and (ND_{DR}), the sequence(s) generated by $\mathscr{F}$ has finite manifold identification property.
3. **Local linearization** There exists a linear matrix $M_{\mathscr{F}}$ such that along the identified smooth manifold(s) the global non-linear iteration locally can be linearized

$$z_{k+1} - z_k = M_{\mathscr{F}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|). \tag{3.1}$$

4. **Spectrum of $M_{\mathscr{F}}$** Owing to the structure of the optimization problem and first-order method, $M_{\mathscr{F}}$ will have certain spectral properties, *e.g.* real or complex spectrum.
5. **Trajectories of $\{z_k\}_{k\in\mathbb{N}}$** The leading eigenvalue of $M_{\mathscr{F}}$ determines the trajectory of $\{z_k\}_{k\in\mathbb{N}}$.

Remark 3.1.

- "Steps 1-4" of the above framework are also the essential steps of the local linear convergence analysis framework for first-order methods [45]. For example, if the spectral radius of $M_{\mathscr{F}}$ is strictly smaller than 1, then one can derive the local linear convergence result.
- The finite manifold identification is not necessarily for $\{z_k\}_{k\in\mathbb{N}}$, as in general first-order methods generate several different points along iteration. Take Douglas–Rachford splitting method (see Eq. (3.3)) for example, $\{z_k\}_{k\in\mathbb{N}}$ is the fixed-point sequence of the method, however the identification is for the shadow sequences $\{u_k\}_{k\in\mathbb{N}}$ and $\{x_k\}_{k\in\mathbb{N}}$; See Section 3.3 for details.
- The o-terms in (3.1) are due to the non-linearity of $\mathscr{F}$ and the curvature of the identified manifold(s). In a series of work [46, 45, 47], the linearization is considered with respect to $z^\star$, that is $z_{k+1} - z^\star = M_{\mathscr{F}}(z_k - z^\star) + o(\|z_k - z^\star\|)$. The main reason of linearization in terms of z_{k+1} and z_k is to better motivate the acceleration scheme in Section 5.

Remark 3.2 (What determines the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ continued). Though the linearization (3.1) makes it possible to analyze the trajectory of $\{z_k\}_{k\in\mathbb{N}}$, it is still quite complicated, as the linearization (3.1) contains a small o-term which is different from Remark 2.2. When $M_{\mathscr{F}}$ contains complex eigenvalues, the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ can only be obtained without the small o-term, which requires the optimization problem to be *locally polyhedral* around the solution.

In the following, we apply the above framework to analyze the trajectory of three classical first-order algorithms: Forward–Backward splitting [51], Douglas–Rachford/ADMM [33, 37] and Primal–Dual splitting [24]. For the purpose of readability, in this section we mainly provide the qualitative description of the trajectory of these methods (*e.g.* which type), and omit the quantitative characterization (*e.g.* speed of convergence of $\cos(\theta_k)$). All the proofs for propositions in this section are provided in Section B.

3.2 Forward–Backward splitting

Forward–Backward splitting [51] is designed to solve the following optimization problem

$$\min_{x\in\mathbb{R}^n} \{\Phi(x) \overset{\text{def}}{=} R(x) + F(x)\}, \tag{$\mathscr{P}_{\text{FB}}$}$$

where the following assumptions are imposed

(F.1) $R \in \Gamma_0(\mathbb{R}^n)$ is proper convex and lower semi-continuous.
(F.2) $F \in C^{1,1}(\mathbb{R}^n)$ is convex differentiable with gradient ∇F being L-Lipschitz continuous.
(F.3) $\text{Argmin}(\Phi) \neq \emptyset$, *i.e.* the set of minimizers is non-empty.

The iteration of Forward–Backward splitting method is described in Algorithm 1.

```
Algorithm 1: Forward–Backward splitting
  Input: $\gamma \in ]0, 2/L[$.
  Initial: $x_0 \in \mathbb{R}^n$.
  Repeat:
      $x_{k+1} = \text{prox}_{\gamma R}\big(x_k - \gamma\nabla F(x_k)\big).$        (3.2)
  Until: $\|x_{k+1} - x_k\| \leq \text{tol}$.
```

The fixed-point formulation of Forward–Backward splitting is quite straightforward, which reads

$$x_{k+1} = \mathscr{F}_{\text{FB}}(x_k) \quad \text{where} \quad \mathscr{F}_{\text{FB}} \overset{\text{def}}{=} \text{prox}_{\gamma R}(\text{Id} - \gamma\nabla F).$$

Remark 3.3. In the literature, various inertial variants of Forward–Backward splitting are proposed, such as inertial Forward–Backward and FISTA [11, 23, 47, 50]. However, these methods will not be covered in this paper as the fixed-point operators of these schemes are not *non-expansive*, and trajectory of the sequence and acceleration for these schemes are much more complicated.

Let $x^\star \in \mathrm{Argmin}(R+F)$ be a global minimizer, we impose the following non-degeneracy condition

$$-\nabla F(x^\star) \in \mathrm{ri}\big(\partial R(x^\star)\big). \tag{$\mathrm{ND}_{\mathrm{FB}}$}$$

We refer to [47] for more detailed discussions about these conditions for the local linear convergence of the general Forward–Backward-type splitting methods.

Remark 3.4. Throughout this section, we impose the non-degeneracy conditions, also for the Douglas–Rachford and Primal–Dual splitting methods, for our analysis. Based on a recent work [35], when the function R is so-called "mirror-stratifiable", condition ($\mathrm{ND}_{\mathrm{FB}}$) can be removed. However, we will not dive into this direction, since it will not affect the conclusion of this section.

We have the following result for the trajectory of $\{x_k\}_{k\in\mathbb{N}}$. Redefine $v_k = x_k - x_{k-1}$.

Theorem 3.5. *For problem ($\mathscr{P}_{\mathrm{FB}}$) and the Forward–Backward splitting method (3.2), suppose that assumptions (**F.1**)-(**F.3**) are true, then $\{x_k\}_{k\in\mathbb{N}}$ converges to a global minimizer $x^\star \in \mathrm{Argmin}(\Phi)$. If, moreover, $R \in \mathrm{PSF}_{x^\star}(\mathscr{M}_{x^\star})$, F is locally C^2 around $x^\star$ and condition ($\mathrm{ND}_{\mathrm{FB}}$) holds, there exists a matrix M_{FB} such that for all k large enough*

$$x_{k+1} - x_k = M_{\mathrm{FB}}(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|).$$

Moreover, we have

(i) *All the eigenvalues of M_{FB} are real and lie in $]-1,1]$.*

(ii) *Let σ_2 be the second largest eigenvalue of M_{FB}. If $\|x_{k+1}-x_k\| \asymp \rho^k$ for some $\rho > \sigma_2$, then the angle θ_k is convergent with $\theta_k \to 0$ and $\{x_k\}_{k\in\mathbb{N}}$ is a Type I sequence.*

Remark 3.6.

- The result holds true for varying but convergent step-size $\gamma_k \in [0, 2/L]$.
- The detailed expression of M_{FB} can be found in Section B. Theorem 3.5 implies that the eventual trajectory of $\{x_k\}_{k\in\mathbb{N}}$ for Forward–Backward is a straight line.
- If there holds R is locally polyhedral around $x^\star$, F is quadratic, then for the linearization we have directly $x_{k+1} - x_k = M_{\mathrm{FB}}(x_k - x_{k-1})$ without the o-terms.

Example 3.1. We consider regularized least square

$$\min_{x\in\mathbb{R}^n} R(x) + \tfrac{1}{2}\|Ax - b\|^2$$

to demonstrate the property of $\{\theta_k\}_{k\in\mathbb{N}}$. Two different cases of R are considered: ℓ_1-norm which is polyhedral and nuclear norm which is not polyhedral. We have $A \in \mathbb{R}^{m\times n}$ and for each cases the settings are

ℓ_1-norm $(m,n) = (48, 128)$, the solution $x^\star$ has 14 non-zero elements.
Nuclear norm $(m,n) = (868, 1024)$, the solution $x^\star$ has rank of 2.

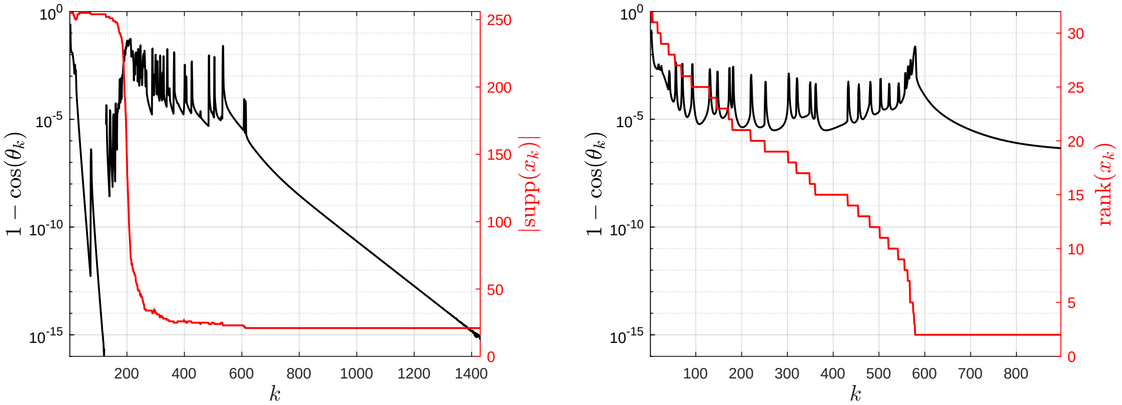
For both examples, A is generated from the standard random Gaussian ensemble. The numerical results are shown in Figure 1. For ℓ_1-norm, besides θ_k, we also provide the change of support size of x_k, *i.e.* $|\mathrm{supp}(x_k)|$:

- For the support of x_k, three phases can be observed: at beginning x_k is almost in the whole space, then the size of supports starts to decrease and eventually becomes stable which is the activity identification.
- The behavior of θ_k also has three phases: 1) when x_k is in the whole space, θ_k is equal or very close to 0; 2) When the support is decreasing, θ_k oscillates; 3) After identification, θ_k converges to 0 linearly.

For nuclear norm, the change of rank of x_k is provided

- Different form the ℓ_1-norm, the rank of x_k gradually decreases, results in a staircase observation.

- For θ_k, inside each staircase, it decrease first and then increases. But after identification of the rank, it converges to 0 linearly.



(a) ℓ_1-norm

(b) Nuclear norm

Figure 1: Finite activity identification and property of θ_k for Forward–Backward splitting method.

3.3 Douglas–Rachford splitting and ADMM

The second example is Douglas–Rachford splitting method [33], for solving the sum of two non-smooth functions

$$\min_{x\in\mathbb{R}^n} R(x)+J(x), \qquad (\mathscr{P}_{\mathrm{DR}})$$

where we assume

(D.1) $R,J\in\Gamma_0(\mathbb{R}^n)$, the proper convex and lower semi-continuous functions.
(D.2) $\mathrm{ri}(\mathrm{dom}(R))\cap\mathrm{ri}(\mathrm{dom}(J))\neq\emptyset$, *i.e.* the domain qualification condition.
(D.3) $\mathrm{Argmin}(R+J)\neq\emptyset$, *i.e.* the set of minimizers is non-empty.

The standard Douglas–Rachford splitting method [33] is described in Algorithm 2.

Algorithm 2: Douglas–Rachford splitting

```
Input: $\gamma > 0$.
Initial: $z_0 \in \mathbb{R}^n$, $x_0 = \mathrm{prox}_{\gamma J}(z_0)$;
Repeat:
    $u_{k+1} = \mathrm{prox}_{\gamma R}(2x_k - z_k),$
    $z_{k+1} = z_k + u_{k+1} - x_k,$          (3.3)
    $x_{k+1} = \mathrm{prox}_{\gamma J}(z_{k+1}),$
Until: $\|z_{k+1} - z_k\| \leq \mathrm{tol}$.
```

The fixed-point formulation of Douglas–Rachford with respect to z_k is

$$z_{k+1}=\mathscr{F}_{\mathrm{DR}}(z_k) \quad \text{where} \quad \mathscr{F}_{\mathrm{DR}} \overset{\text{def}}{=} \tfrac{1}{2}\big((2\mathrm{prox}_{\gamma R}-\mathrm{Id})(2\mathrm{prox}_{\gamma J}-\mathrm{Id})+\mathrm{Id}\big).$$

Remark 3.7. It is well known that the alternating direction method of multipliers (ADMM) is closely connected with Douglas–Rachford splitting method, for its local trajectory property, we refer to [66] for detailed discussion.

Below we first present the linearization of Douglas–Rachford iteration (3.3) and then discuss the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ under two different cases: both R,J in ($\mathscr{P}_{\mathrm{DR}}$) are non-smooth as in (**D.1**), and one of the functions is smooth. We shall see that two different trajectories are exhibited by the method.

3.3.1 Linearization of Douglas–Rachford splitting

Let $z^\star \in \mathrm{fix}(\mathscr{F}_{\mathrm{DR}})$ and $x^\star = \mathrm{prox}_{\gamma J}(z^\star) \in \mathrm{Argmin}(R+J)$ such that $z_k \to z^\star$ and $x_k, u_k \to x^\star$, from (3.3) the corresponding first-order optimality condition reads $x^\star - z^\star \in \gamma\partial R(x^\star)$ and $z^\star - x^\star \in \gamma\partial J(x^\star)$. We assume the following non-degeneracy condition

$$x^\star - z^\star \in \gamma\mathrm{ri}\big(\partial R(x^\star)\big) \quad \text{and} \quad z^\star - x^\star \in \gamma\mathrm{ri}\big(\partial J(x^\star)\big). \tag{ND$_{\mathrm{DR}}$}$$

Theorem 3.8. *For problem ($\mathscr{P}_{\mathrm{DR}}$) and the Douglas–Rachford splitting algorithm (3.3), suppose that the conditions* (**D.1**)-(**D.3**) *are true, then $\{z_k\}_{k\in\mathbb{N}}$ converges to a point $z^\star \in \mathrm{fix}(\mathscr{F}_{\mathrm{DR}})$ and $\{x_k\}_{k\in\mathbb{N}}, \{u_k\}_{k\in\mathbb{N}}$ converge to $x^\star \overset{\mathrm{def}}{=} \mathrm{prox}_{\gamma R}(z^\star) \in \mathrm{Argmin}(R+J)$. If moreover, $R \in \mathrm{PSF}_{x^\star}(\mathscr{M}^R_{x^\star})$ and $J \in \mathrm{PSF}_{x^\star}(\mathscr{M}^J_{x^\star})$ are partly smooth and condition* (ND$_{\mathrm{DR}}$) *holds, then there exists a matrix M_{DR} such that for all k large enough*

$$z_{k+1} - z_k = M_{\mathrm{DR}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|).$$

See Section B.3.2 for the proof and expression of M_{DR}. We refer to [48] for detailed discussions on the local linear convergence of Douglas–Rachford splitting method.

3.3.2 Trajectory of Douglas–Rachford splitting

We first consider the case that both R, J are non-smooth. Let $R \in \mathrm{PSF}_{x^\star}(\mathscr{M}^R_{x^\star}), J \in \mathrm{PSF}_{x^\star}(\mathscr{M}^J_{x^\star})$, and denote $T^R_{x^\star}, T^J_{x^\star}$ the tangent spaces of $\mathscr{M}^R_{x^\star}, \mathscr{M}^J_{x^\star}$ at $x^\star$, respectively. Denote θ_F the Friedrichs angle between $T^R_{x^\star}$ and $T^J_{x^\star}$.

Theorem 3.9. *For problem ($\mathscr{P}_{\mathrm{DR}}$) and the Douglas–Rachford splitting algorithm iteration (3.3), assume that Theorem 3.8 holds. If, moreover, R, J are locally polyhedral around $x^\star$, then $z_{k+1} - z_k = M_{\mathrm{DR}}(z_k - z_{k-1})$ with*

$$M_{\mathrm{DR}} = \mathcal{P}_{T^R_{x^\star}}\mathcal{P}_{T^J_{x^\star}} + (\mathrm{Id} - \mathcal{P}_{T^R_{x^\star}})(\mathrm{Id} - \mathcal{P}_{T^J_{x^\star}}).$$

If moreover $\dim(T^R_{x^\star} \cap T^J_{x^\star}) < \min\big\{\dim(T^R_{x^\star}), \dim(T^J_{x^\star})\big\}$, the angle θ_k is convergent to $\theta_F \in]0, \pi/2]$ and $\{z_k\}_{k\in\mathbb{N}}$ is a Type II sequence.

Remark 3.10.

- The spectral properties of M_{DR} is much more difficult to analyze compared to that of Forward–Backward splitting, and for the case both R, J are non-smooth, local polyhedrality around the solution is need. For this setting, M_{DR} is a normal matrix, hence quasi-diagonalizable [40, Theorem 2.5.8], with the leading block of the decomposition reads

  $$B = \cos(\theta_F)\begin{bmatrix} \cos(\theta_F) & \sin(\theta_F) \\ -\sin(\theta_F) & \cos(\theta_F) \end{bmatrix}.$$

  Condition $\dim(T^R_{x^\star} \cap T^J_{x^\star}) < \min\big\{\dim(T^R_{x^\star}), \dim(T^J_{x^\star})\big\}$ ensures that $\theta_F \in]0, \pi/2]$ which makes B a rotation. Consequently the local trajectory of the sequence $\{z_k\}_{k\in\mathbb{N}}$ is a *logarithmic spiral*. Moreover, the choice of γ does not affect the local trajectory of $\{z_k\}_{k\in\mathbb{N}}$ as B only depends on the Friedrichs angle θ_F.
- The analysis of the $\cos(\theta_k)$ depends on the explicit expression of the leading eigenvalues of M_{DR}, which is only available when R, J are locally polyhedral around the solution. For the case that R, J are general partly smooth function, the behavior of θ_k depends on γ due to the non-trivial Riemannian Hessian of R, J; See Figure 2.

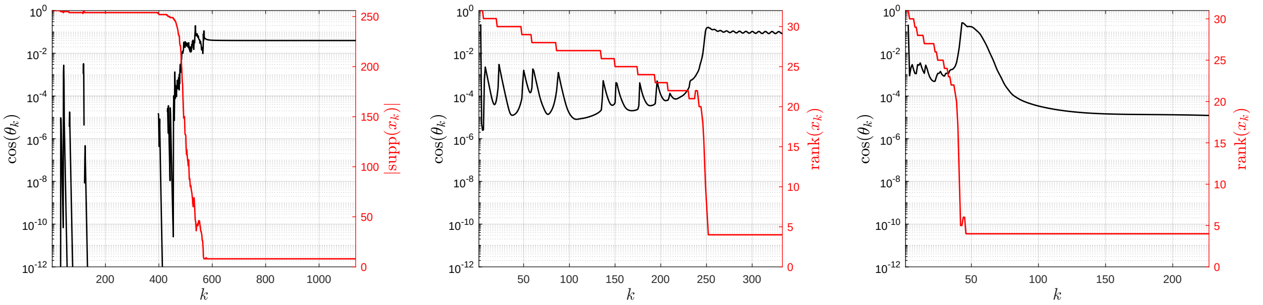
Example 3.2. We use the affine constrained problem

$$\min_{x\in\mathbb{R}^n} R(x) \ \ \text{such that} \ \ Ax = A\mathring{x} \tag{3.4}$$

to demonstrate the property of $\{\theta_k\}_{k\in\mathbb{N}}$. Similar to Example 3.1, ℓ_1-norm and nuclear norm are considered for R, $A \in \mathbb{R}^{m\times n}$ is generated from the standard random Gaussian ensemble and

ℓ_1-norm $(m,n) = (48, 128)$, $\mathring{x}$ has 8 non-zero elements.
Nuclear norm $(m,n) = (620, 1024)$, $\mathring{x}$ has rank of 2.

The results are shown in Figure 2, the observations of ℓ_1-norm are similar to those in Example 3.1, except that θ_k eventually converges to some non-zero values. For nuclear norm, two choices of γ, $\gamma = 1, 6$, are considered. Observe that after rank identification, θ_k oscillates in an interval for $\gamma = 1$ and behaves smoothly for $\gamma = 6$.



(a) ℓ_1-norm

(b) Nuclear norm: $\gamma = 1$

(c) Nuclear norm: $\gamma = 6$

Figure 2: Finite activity identification and property of θ_k for Douglas–Rachford splitting method for solving affine constrained problem.

Remark 3.11. As nuclear norm is not polyhedral, it has non-trivial Riemannian Hessian matrix. Therefore, the choices of γ affects the eventually behavior of $\{\theta_k\}_{k\in\mathbb{N}}$. While for ℓ_1-norm, the value that $\{\theta_k\}_{k\in\mathbb{N}}$ converges to is independent of γ.

Assume now R is locally C^2-smooth around the solution, we shall see that different from the above polyhedral case, the choice of γ will impact the trajectory of $\{z_k\}_{k\in\mathbb{N}}$.

Theorem 3.12. *For problem ($\mathscr{P}_{\mathrm{DR}}$) and the Douglas–Rachford splitting algorithm (3.3), assume conditions (**D.1**)-(**D.3**) are true. If R is locally C^2 around $x^\star$ and $J \in \mathrm{PSF}_{x^\star}(\mathcal{M}^J_{x^\star})$ is partly smooth and condition ($\mathrm{ND}_{\mathrm{DR}}$) holds for J, then Theorem 3.8 holds. Moreover,*

(i) *All the eigenvalues of M_{DR} are real if γ is chosen such that $\gamma < \frac{1}{\|\nabla^2 R(x^\star)\|}$.*

(ii) *For the angle θ_k, we have $\theta_k \to 0$, and $\{z_k\}_{k\in\mathbb{N}}$ is a Type I sequence.*

Remark 3.13.

- The result also holds true for the case when both R, J are smooth.
- When $\gamma \geq \frac{1}{\|\nabla^2 R(x^\star)\|}$, M_{DR} will have complex eigenvalues, however not necessarily for the leading eigenvalue, as a result the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ can be either straight line or spiral.

We refer to Figure 5 for example of applying Douglas–Rachford splitting method to solve LASSO problem, on how the choice of γ affects the trajectory of $\{z_k\}_{k\in\mathbb{N}}$.

3.4 Primal–Dual splitting

For problem ($\mathscr{P}_{\mathrm{DR}}$), consider function J is composed with a linear mapping L

$$\min_{x\in\mathbb{R}^n} R(x) + J(Lx), \tag{$\mathscr{P}_{\mathrm{PD}}$}$$

where we assume

(**P.1**) $R \in \Gamma_0(\mathbb{R}^n)$ and $J \in \Gamma_0(\mathbb{R}^m)$.
(**P.2**) $L : \mathbb{R}^n \to \mathbb{R}^m$ is a linear mapping.
(**P.3**) The inclusion $0 \in \mathrm{ran}(\partial R + L^T \circ \partial J \circ L)$ holds.

The problem above can be handled efficiently by ADMM, in the literature, another popular approach is the Primal–Dual splitting method. The saddle-point problem associated to ($\mathscr{P}_{\mathrm{PD}}$) reads

$$\min_{x\in\mathbb{R}^n} \max_{w\in\mathbb{R}^m} R(x) + \langle Lx,\, w\rangle - J^*(w), \tag{$\mathscr{P}_{\mathrm{SP}}$}$$

where J^* is the Legendre-Fenchel conjugate of J. If we fully dualize ($\mathscr{P}_{\mathrm{PD}}$), then we obtain its Fenchel-Rockafellar dual form

$$\min_{w \in \mathbb{R}^m} R^*(-L^T w) + J^*(w). \tag{$\mathscr{D}_{\mathrm{PD}}$}$$

Denote by $\mathscr{X}$ and $\mathscr{W}$ the sets of solutions of problem ($\mathscr{P}_{\mathrm{PD}}$) and ($\mathscr{D}_{\mathrm{PD}}$), respectively.

Below we describe a Primal–Dual splitting method [24] for solving the saddle point problem.

```
Algorithm 3: A Primal–Dual splitting method
  Input: $\gamma_R, \gamma_J > 0$ such that $\gamma_R \gamma_J \|L\|^2 < 1$ and $\tau \in [0,1]$.
  Initial: $x_0 \in \mathbb{R}^n$, $w_0 \in \mathbb{R}^m$;
  Repeat:
      $x_{k+1} = \mathrm{prox}_{\gamma_R R}(x_k - \gamma_R L^T w_k)$,
      $\bar{x}_{k+1} = x_{k+1} + \tau(x_{k+1} - x_k)$,                    (3.5)
      $w_{k+1} = \mathrm{prox}_{\gamma_J J^*}(w_k + \gamma_J L \bar{x}_{k+1})$,
  Until: $\|x_{k+1} - x_k\| + \|w_{k+1} - w_k\| \leq \mathrm{tol}$.
```

Define the following augmented variable z_k and operators

$$z_k \overset{\text{def}}{=} \begin{pmatrix} x_k \\ w_k \end{pmatrix}, \quad \boldsymbol{A} \overset{\text{def}}{=} \begin{bmatrix} A & L^* \\ -L & C^{-1} \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mathcal{V}} \overset{\text{def}}{=} \begin{bmatrix} \mathrm{Id}_n/\gamma_R & -L^* \\ -L & \mathrm{Id}_m/\gamma_J \end{bmatrix}, \tag{3.6}$$

where $\mathrm{Id}_n, \mathrm{Id}_m$ are the identity operators on $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. We have $\boldsymbol{A}$ is maximal monotone [20] and $\boldsymbol{\mathcal{V}}$ is self-adjoint and ν-positive definite for $\nu = (1 - \sqrt{\gamma_R \gamma_J \|L\|^2}) \min\{\frac{1}{\gamma_R}, \frac{1}{\gamma_J}\}$ [78, 28]. The fixed-point characterization of (3.5) when $\tau = 1$ reads

$$z_{k+1} = (\boldsymbol{\mathcal{V}} + \boldsymbol{A})^{-1} \boldsymbol{\mathcal{V}}(z_k) = (\mathbf{Id} + \boldsymbol{\mathcal{V}}^{-1} \boldsymbol{A})^{-1}(z_k), \tag{3.7}$$

which is a special case of proximal point algorithm [28, 27]. We also refer to [78, 27] for more general form of Primal–Dual splitting methods.

3.4.1 Linearization of Primal–Dual splitting

Let $(x^\star, w^\star) \in \mathscr{X} \times \mathscr{W}$ be a saddle-point, the first-order optimality condition entails $-L^T w^\star \in \partial R(x^\star)$ and $L x^\star \in \partial J^*(w^\star)$. We impose the following non-degeneracy condition

$$-L^T w^\star \in \mathrm{ri}\big(\partial R(x^\star)\big) \quad \text{and} \quad L x^\star \in \mathrm{ri}\big(\partial J^*(w^\star)\big). \tag{$\mathrm{ND}_{\mathrm{PD}}$}$$

Theorem 3.14. *For problem ($\mathscr{P}_{\mathrm{PD}}$) and the Primal–Dual splitting algorithm (3.5), suppose assumptions (**P.1**)-(**P.3**) are true. If $\tau = 1$ and γ_R, γ_J are chosen such that $\gamma_R \gamma_J \|L\|^2 < 1$, then $(x_k, w_k) \to (x^\star, w^\star) \in \mathscr{X} \times \mathscr{W}$. If moreover, $R \in \mathrm{PSF}_{x^\star}(\mathscr{M}^R_{x^\star})$ and $J^* \in \mathrm{PSF}_{w^\star}(\mathscr{M}^{J^*}_{w^\star})$ are partly smooth and condition ($\mathrm{ND}_{\mathrm{PD}}$) holds, then for all k large enough there exists a matrix M_{PD} such that*

$$z_{k+1} - z_k = M_{\mathrm{PD}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|).$$

See Section B.3.3 for the proof and expression of M_{PD}. We refer to [49] for detailed discussions on the local linear convergence of a class of Primal–Dual splitting methods.

3.4.2 Trajectory of Primal–Dual splitting

The trajectory of Primal–Dual splitting also depends on the explicit analysis of the spectrum of M_{PD} which is only available when R, J^* are locally polyhedral around the saddle point. Let $R \in \mathrm{PSF}_{x^\star}(\mathscr{M}^R_{x^\star}), J \in \mathrm{PSF}_{w^\star}(\mathscr{M}^{J^*}_{w^\star})$, and denote $T^R_{x^\star}, T^{J^*}_{w^\star}$ the tangent spaces of $\mathscr{M}^R_{x^\star}, \mathscr{M}^{J^*}_{w^\star}$ at $x^\star$ and $w^\star$, respectively. Denote $\bar{L} \overset{\text{def}}{=} \mathcal{P}_{T^{J^*}_{w^\star}} L \mathcal{P}_{T^R_{x^\star}}$.

Theorem 3.15. *For problem ($\mathscr{P}_{\mathrm{PD}}$) and the Primal–Dual iteration (3.5), assume Theorem 3.14 holds. If, moreover, R, J^* locally are polyhedral around $(x^\star, w^\star)$, then $z_{k+1} - z_k = M_{\mathrm{PD}}(z_k - z_{k-1})$ with*

$$M_{\mathrm{PD}} = \begin{bmatrix} \mathrm{Id}_n & -\gamma_R \bar{L}^T \\ \gamma_J \bar{L} & \mathrm{Id}_m - (1+\tau)\gamma_J\gamma_R \bar{L}\bar{L}^T \end{bmatrix}.$$

Moreover, M_{PD} is block diagonalizable with the leading block being 2×2 which corresponds to elliptical rotation. Then there exist $\underline{\theta}, \overline{\theta}$ such that eventually $\theta_k \in [\underline{\theta}, \overline{\theta}]$, and $\{z_k\}_{k\in\mathbb{N}}$ is a Type III sequence.

Remark 3.16.

- Let σ be the leading eigenvalue of $\bar{L}\bar{L}^T$, then the leading block of the decomposition of M_{PD} reads
  $$B = \begin{bmatrix} 1 & -\gamma_R\sigma \\ \gamma_J\sigma & 1-(1+\tau)\gamma_R\gamma_J\sigma^2 \end{bmatrix}.$$
  Owing to Proposition A.11, there exist some $\psi, \phi \in [0, \pi/2]$ and $l, s > 0$ such that
  $$B = \frac{1}{\sqrt{1-\tau\gamma_J\gamma_R\sigma^2}} \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix} \begin{bmatrix} \cos(\phi) & \frac{s}{l}\sin(\phi) \\ -\frac{l}{s}\sin(\phi) & \cos(\phi) \end{bmatrix},$$
  with
  $$\begin{aligned} \psi &= \mathrm{arccot}\Big(-\tfrac{\gamma_J - \gamma_R}{(1+\tau)\gamma_R\gamma_J\sigma}\Big), \\ \phi &= \arccos\Big(\tfrac{(\gamma_R+\gamma_J)\sigma\sin(\psi)+(2-(1+\tau)\gamma_R\gamma_J\sigma^2)\cos(\psi)}{2\sigma}\Big) \\ \tfrac{s}{l} &= \tfrac{1-(1+\tau)\gamma_R\gamma_J\sigma^2}{\sin(\psi)\sin(\phi)\sigma} - \cot(\psi)\cot(\phi). \end{aligned}$$
  This means that B is a composition of circular rotation and elliptical rotation discussed in Proposition A.9. Therefore, invoke the result of Proposition A.7 and A.9, we have $\underline{\theta} = \psi - \overline{\chi}, \overline{\theta} = \psi - \underline{\chi}$ with
  $$\cos(\overline{\chi}) = \frac{(\frac{s}{2l}+\frac{l}{2s})\cos(\phi)-|\frac{s}{2l}-\frac{l}{2s}|}{\sqrt{\sin^2(\phi)+((\frac{s}{2l}+\frac{l}{2s})\cos(\phi)-|\frac{s}{2l}-\frac{l}{2s}|)^2}} \quad \text{and} \quad \cos(\underline{\chi}) = \frac{(\frac{s}{2l}+\frac{l}{2s})\cos(\phi)+|\frac{s}{2l}-\frac{l}{2s}|}{\sqrt{\sin^2(\phi)+((\frac{s}{2l}+\frac{l}{2s})\cos(\phi)+|\frac{s}{2l}-\frac{l}{2s}|)^2}},$$
- Similar to the case of Douglas–Rachford splitting, the trajectory of Primal–Dual, when both R, J^* are locally polyhedral, is obtained via the explicit analysis of the spectrum of M_{PD} which will not be available when R, J^* are general partly smooth functions. Moreover, unlike the case of Douglas–Rachford, when one of R, J^* is locally C^2-smooth, the trajectory of Primal–Dual splitting remains a spiral.

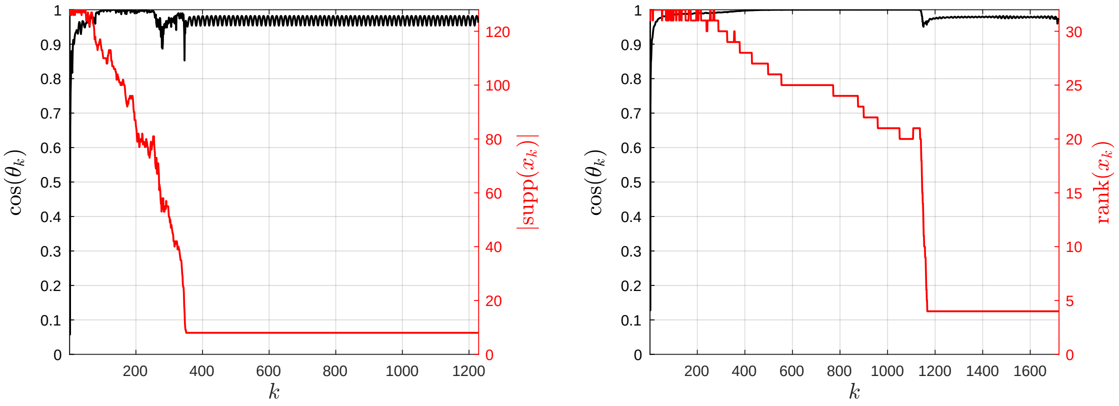
Example 3.3. We continue using the problem (3.4) in Example 3.2 to demonstrate the trajectories of Primal–Dual splitting method. The observations are shown in Figure 3,

- For ℓ_1-norm, θ_k eventually oscillates in an interval which complies with our result in Theorem 3.15.
- For nuclear norm, though it is not covered by our result as nuclear norm is not polyhedral, locally the value of θ_k also oscillates.

4 The failure of inertial technique

The trajectory results from previous section provide a geometric explanation why inertial acceleration works for (proximal) gradient descent methods but not the others. For (proximal) gradient descent, as the trajectory of the generated sequence eventually approximates a straight line, the direction of $x_k - x_{k-1}$ points towards the solution, hence moving certain distance along the inertial direction provides acceleration. However, when the trajectory of the generated sequence is a spiral, as for the cases of Douglas–Rachford and Primal–Dual splitting, the direction of $z_k - z_{k-1}$ does not point toward the solution, hence fail to provide acceleration.

In this section, we consider Douglas–Rachford and two different problems to demonstrate the outcomes of inertial acceleration. We show that the performance of inertial Douglas–Rachford is both problem and parameter dependent. Specialize the inertial scheme (1.3) to the case of Douglas–Rachford splitting, we obtain an inertial Douglas–Rachford splitting scheme described in Algorithm 4.



(a) ℓ_1-norm

(b) Nuclear norm

Figure 3: Finite activity identification and property of θ_k for Primal–Dual splitting method for solving affine constrained problem.

```
Algorithm 4: An inertial Douglas–Rachford splitting
    Input: $\gamma > 0$.
    Initial: $z_0 \in \mathbb{R}^n$, $\bar{z}_0 = z_0$, $x_0 = \mathrm{prox}_{\gamma J}(\bar{z}_0)$;
    Repeat:
        $u_{k+1} = \mathrm{prox}_{\gamma R}(2x_k - \bar{z}_k)$,
        $z_{k+1} = \bar{z}_k + u_{k+1} - x_k$,
        $\bar{z}_{k+1} = z_{k+1} + a_k(z_{k+1} - z_k)$,        (4.1)
        $x_{k+1} = \mathrm{prox}_{\gamma J}(\bar{z}_{k+1})$,
    Until: $\|z_{k+1} - z_k\| \leq \mathrm{tol}$.
```

4.1 Feasibility problem

We first consider a feasibility problem of two subspaces. For simplicity, consider the problem in $\mathbb{R}^2$: let $T_1, T_2 \subset \mathbb{R}^2$ be two intersecting lines. The problem of finding the common point of T_1, T_2 can be written as

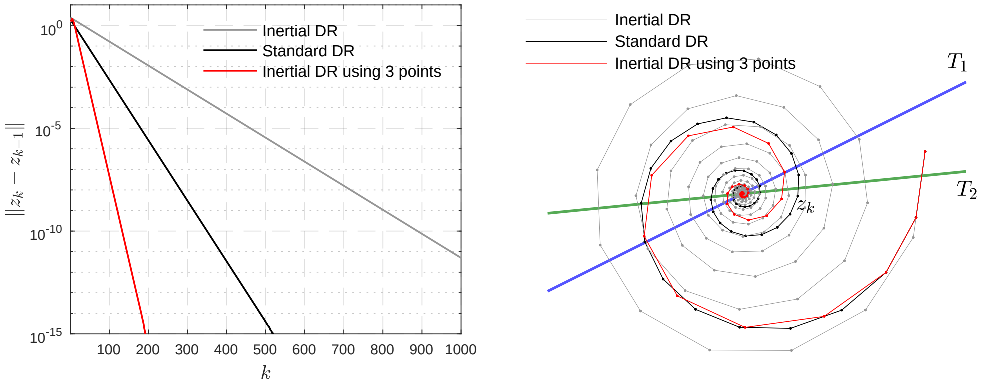
$$\min_{x\in\mathbb{R}^2} \iota_{T_1}(x) + \iota_{T_2}(x). \tag{4.2}$$

As the proximal mapping of indicator functions is projection, the above problem can be easily handle by Douglas–Rachford splitting method.

For the inertial Douglas–Rachford (4.1), we consider $a_k \equiv 0.3$ and compare it with the standard Douglas–Rachford splitting scheme (3.3). The comparison is provided in Figure 4, with the left figure showing the convergence speed of $\|z_k - z_{k-1}\|$ and right figure the trajectory of sequence $\{z_k\}_{k\in\mathbb{N}}$. We observe that

- The inertial Douglas–Rachford with $a_k \equiv 0.3$ (*gray* line) is slower than the standard scheme (*black* line). Moreover, it can be shown that for this feasibility example, the inertial Douglas–Rachford is slower as long as $a_k > 0$; See Section A of [66].
- The slow performance of inertial Douglas–Rachford can also be visualized by the trajectory of the sequence $\{z_k\}_{k\in\mathbb{N}}$. For inertial Douglas–Rachford, it increases the length of the trajectory of $\{z_k\}_{k\in\mathbb{N}}$, see the difference between *gray* and *black* spirals.

The above comparisons, is quite different from the improvement of *e.g.* heavy-ball method over gradient descent, which is an evidence that the trajectory of the sequence affects the outcome of inertial acceleration. We also remark that in [45, Chapter 4], the author suggested to use more than two points for computing $\bar{z}_k$ in



(a) Convergence of $\|z_k - z_{k-1}\|$ (b) Trajectory of $\{z_k\}_{k\in\mathbb{N}}$

Figure 4: Comparison between standard Douglas–Rachford and inertial Douglas–Rachford. (a) convergence of $\|z_k - z_{k-1}\|$; (b) trajectory of sequence $\{z_k\}_{k\in\mathbb{N}}$.

(4.1), for example the following three-point approach

$$\bar{z}_k = z_k + a_k(z_k - z_{k-1}) + b_k(z_{k-1} - z_{k-2}). \tag{4.3}$$

Particularly, with $a_k > 0$ and $b_k < 0$, the above choice can provide acceleration. In Figure 4, we also provide such a test with $(a_k, b_k) = (0.6, -0.3)$, the corresponding convergence observation of $\|z_k - z_{k-1}\|$ is shown in *red* line which is faster than the standard Douglas–Rachford splitting scheme. Trajectory wise, the length is also shorter than that of the standard Douglas–Rachford.

Remark 4.1.

- The observation is not limited to the simple feasibility case, but rather a class of problems. For example, as long as the problem is (locally) polyhedral around the solution, the above observation can be obtained.
- The difference between the inertial scheme (4.1) and (4.3) implies that when the trajectory is a spiral, only $z_k - z_{k-1}$ is not enough to estimate or fit the direction that z_k travels, while the three-point scheme can solve the problem. However, the problem with the three-point scheme is that, the value of b_k needs to be negative, and in general there is no good way to determine the choices of a_k, b_k, let alone theoretical acceleration guarantees. This is one motivation behind the adaptive acceleration in Section 5.

4.2 LASSO problem

The second problem we consider is the LASSO problem

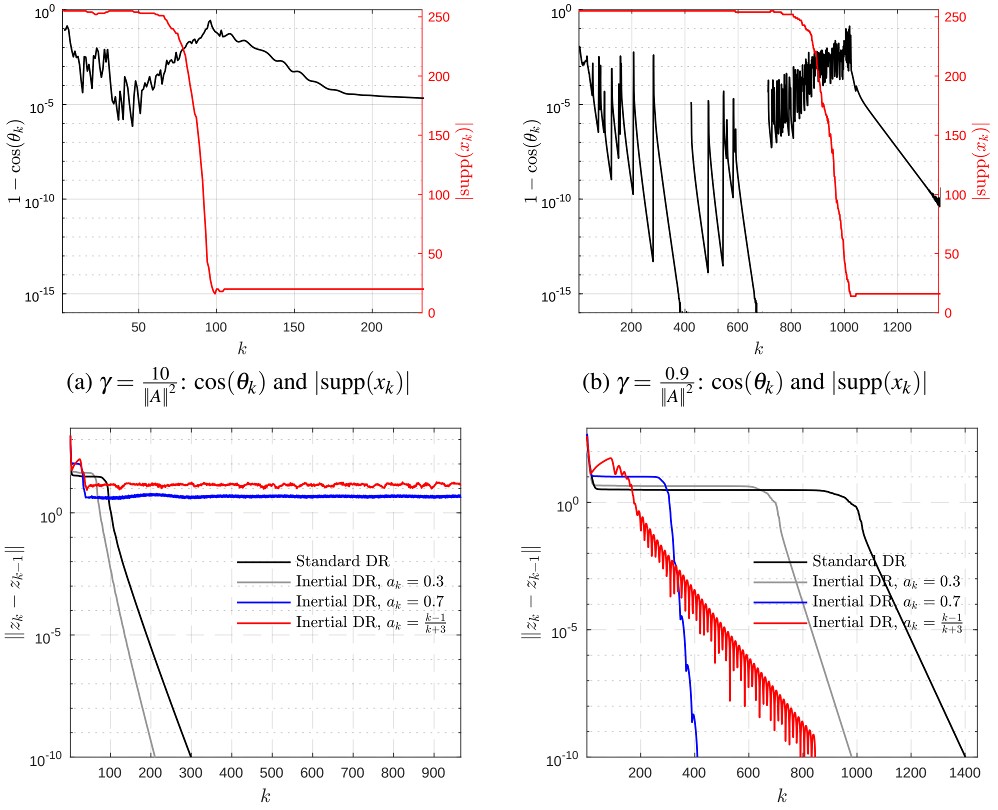
$$\min_{x\in\mathbb{R}^n} \mu\|x\|_1 + \tfrac{1}{2}\|Ax - f\|^2, \tag{4.4}$$

where $A \in \mathbb{R}^{m\times n}$ is random Gaussian matrix with $m < n$.

Since $\frac{1}{2}\|Ax - b\|^2$ is C^2 smooth, we know from Theorem 3.12 that the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is determined by the choice of γ. As a result, two different choices of γ, $\gamma \in \{\frac{0.9}{\|A\|^2}, \frac{10}{\|A\|^2}\}$, are considered. For each γ, four different choices of a_k are chosen: $a_k \equiv 0, a_k \equiv 0.3, a_k \equiv 0.7$ and $a_k = \frac{k-1}{k+3}$. Note that the last choice of a_k corresponds to the Nesterov's scheme of [76] and the FISTA scheme of [23].

For the numerical example, we consider $K \in \mathbb{R}^{64\times 256}$ and $\mu = 2$, f is the measurement of an $\mathring{x}$ which is 8-sparse under small additive white Gaussian noise. The results are shown in Figure 5,

- Case $\gamma = \frac{10}{\|A\|^2}$: in Figure 5 (a), the *red* line shows the support identification of the iterates x_k, and after support identification which is about $k = 150$, the angle θ_k is *not* converging to 0.
- Case $\gamma = \frac{0.9}{\|A\|^2}$: In Figure 5 (b), from about $k = 1600$, θ_k is converging to 0.



(a) $\gamma = \frac{10}{\|A\|^2}$: $\cos(\theta_k)$ and $|\mathrm{supp}(x_k)|$

(b) $\gamma = \frac{0.9}{\|A\|^2}$: $\cos(\theta_k)$ and $|\mathrm{supp}(x_k)|$

(c) $\gamma = \frac{10}{\|A\|^2}$: Convergence of $\|z_k - z_{k-1}\|$

(d) $\gamma = \frac{0.9}{\|A\|^2}$: Convergence of $\|z_k - z_{k-1}\|$

Figure 5: Comparison between standard Douglas–Rachford and inertial Douglas–Rachford. (a) $\cos(\theta_k)$ and $\mathrm{supp}(x_k)$ for $\gamma = \frac{10}{\|K\|^2}$; (b) $\cos(\theta_k)$ and $\mathrm{supp}(x_k)$ for $\gamma = \frac{0.9}{\|K\|^2}$; (c) convergence of $\|z_k - z_{k-1}\|$ for $\gamma = \frac{10}{\|K\|^2}$; (b) trajectory of sequence $\{z_k\}_{k\in\mathbb{N}}$ for $\gamma = \frac{0.9}{\|K\|^2}$.

Then in terms of the performances of the inertial schemes,

- Case $\gamma = \frac{10}{\|A\|^2}$: in Figure 5 (c), out of three inertial schemes, only the one with $a_k \equiv 0.3$ is convergent. This is due to that fact that the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ is a spiral for $\gamma = \frac{10}{\|A\|^2}$.
- Case $\gamma = \frac{0.9}{\|A\|^2}$: All choices of a_k work since $\{z_k\}_{k\in\mathbb{N}}$ eventually forms a straight line. Among these four choices of a_k, $a_k = 0.7$ is the fastest, with the FISTA choice a bit slower.

It can be observed that, under the two choices of γ, the standard Douglas–Rachford is faster for $\gamma = \frac{10}{\|A\|^2}$ than $\gamma = \frac{0.9}{\|A\|^2}$. However, we remark that our main focus here is to demonstrate how the trajectory of $\{z_k\}_{k\in\mathbb{N}}$ affects the outcome of inertial acceleration.

4.3 A geometric interpretation on the failure of inertial

In this part, we provide a geometric explanation on how the trajectory of sequence affects the outcome of inertial acceleration, similar analysis can also be obtained for over-relaxation. To this end, consider the angle ϑ_k defined by $\vartheta_k \overset{\text{def}}{=} \angle(z_k - z_{k-1}, z^\star - z_k) = \arccos\big(\frac{\langle z_k - z_{k-1}, z^\star - z_k\rangle}{\|z_k - z_{k-1}\|\|z^\star - z_k\|}\big)$. The motivation of considering this angle is shown below in Figure 6: let $\bar{z}_k = z_k + a_k(z_k - z_{k-1})$ with $a_k > 0$

- If ϑ_k is *acute*, then it can be shown that there holds $\|z^\star - \bar{z}_k\| < \|z^\star - z_k\|$ as long as

$$a_k < \frac{2\cos(\vartheta_k)\|z^\star - z_k\|}{\|z_k - z_{k-1}\|}.$$

When $a_k = \frac{\cos(\vartheta_k)\|z^\star - z_k\|}{\|z_k - z_{k-1}\|}$, $\bar{z}_k$ is the closest point to $z^\star$ in the radial line of $z_k - z_{k-1}$.

- If ϑ_k is right or *obtuse*, from the above equation we get that $\|z^\star - \bar{z}_k\| < \|z^\star - z_k\|$ holds only for non-positive z_k, which means extrapolate only slows down the convergence.

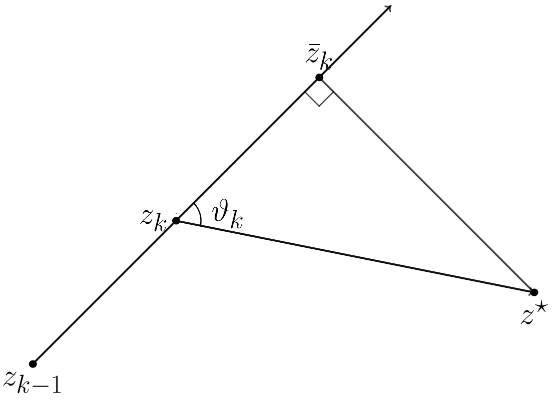


Figure 6: Illustration between inertial direction and the direction to the limiting point in $\mathbb{R}^2$.

In the following, we consider the following linear system in $\mathbb{R}^2$:

$$z_k = Mz_{k-1}$$

with z_k converging to some $z^\star$, and study the property of ϑ_k, under different M that corresponds to the three types of trajectories.

Straight line trajectory Let U be a unitary 2×2 matrix and σ_1, σ_2 be such that $0 < |\sigma_2| < \sigma_1 < 1$, and let M of the form

$$M = U \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} U^T.$$

It is immediate that $z^\star = 0$. Denote $y_k = U^T z_k$, we have

$$y_k = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} y_{k-1} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}^k y_0 = \begin{bmatrix} \sigma_1 & 0 \\ 0 & 0 \end{bmatrix}^k y_0 + \begin{bmatrix} 0 & 0 \\ 0 & \sigma_2 \end{bmatrix}^k y_0 = \sigma_1^k \begin{pmatrix} a \\ \eta^k b \end{pmatrix},$$

where $\eta = \sigma_2/\sigma_1 < 1$. Assume that $y_0 = \begin{pmatrix} a \\ b \end{pmatrix}$, then

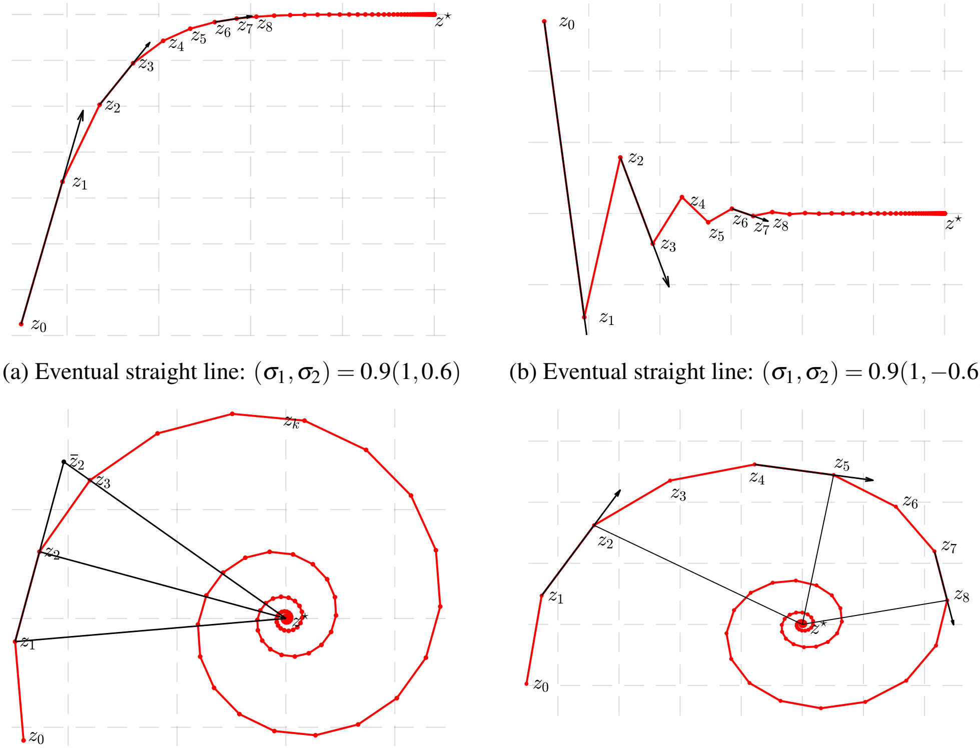
$$\cos(\vartheta_k) = \frac{\langle z_k - z_{k-1}, -z_k \rangle}{\|z_k - z_{k-1}\|\|-z_k\|} = \frac{\langle y_k - y_{k-1}, -y_k \rangle}{\|y_k - y_{k-1}\|\|-y_k\|} = \frac{(1-\sigma_1)a^2 + (1-\sigma_1\eta)\sigma_1\eta^{2k-1}b^2}{\sqrt{(\sigma_1-1)^2a^2 + (\sigma_1\eta-1)^2\eta^{2(k-1)}}\sqrt{a^2+\sigma_1^2\eta^{2k}}}.$$

Let $k \to +\infty$ we get $\cos(\vartheta_k) \to 1$ which means $\vartheta_k \to 0$.

The above result implies that, eventually $z_k - z_{k-1}$ points towards the limiting point $z^\star$. Therefore, moving certain distance along $z_k - z_{k-1}$ is useful to improve the convergence speed. To demonstrate the above result, we consider two different choices of (σ_1, σ_2) that $(\sigma_1, \sigma_2) = 0.9(1, \pm 0.6)$. The trajectory of $\{z_k\}_{k\in\mathbb{N}}$ and the property of ϑ_k are shown in Figure 7 (a) and (b). It can be observed from both figures that: from the beginning, $z_k - z_{k-1}$ is not pointing towards $z^\star$ but eventually almost directly to $z^\star$.

Logarithmic spiral trajectory For logarithmic spiral, we can show that inertial always slows down the convergence. For this case, we have M of the form

$$M = \cos(\theta) \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix},$$



(a) Eventual straight line: $(\sigma_1, \sigma_2) = 0.9(1, 0.6)$

(b) Eventual straight line: $(\sigma_1, \sigma_2) = 0.9(1, -0.6)$

(c) Logarithmic spiral

(d) Elliptical spiral

Figure 7: Graphical illustration of the direction of $z_k - z_{k-1}$ for different types of sequence trajectory.

for some $\theta \in]0, \pi/2[$. For the sequence $\{z_k\}_{k\in\mathbb{N}}$, we also have $z_k \to 0$ as $\rho(M) = \cos(\theta) < 1$. Given any k, we have $\|z_k\| = \|Mz_{k-1}\| = \|z_{k-1}\|\cos(\theta)$, then consider the inner product

$$\begin{aligned}\langle z_k - z_{k-1},\, z_k\rangle = \|z_k\|^2 - \langle z_{k-1},\, z_k\rangle &= \|z_{k-1}\|^2\cos^2(\theta) - \|z_{k-1}\|\|z_k\|\cos(\theta)\\ &= \|z_{k-1}\|^2\cos^2(\theta) - \|z_{k-1}\|^2\cos^2(\theta) = 0,\end{aligned}$$

which means $\cos(\vartheta_k) \equiv 0$ and $\vartheta_k \equiv \pi/2$.

A graphic illustration is provided in Figure 7 (c). Let $k = 2$, and $\bar{z}_2 = 2z_2 - z_1$. It can be proved that the three points $z_1, \bar{z}_2$ and $z^\star$ form an isosceles triangle with $\|z_1 - z^\star\| = \|\bar{z}_2 - z^\star\|$. Moreover, $\bar{z}_2, z_3$ and $z^\star$ are in the same line. This in turn indicates that for all the point z in the segment of z_2 and $\bar{z}_2$, we have

$$\|z_2 - z^\star\| < \|z - z^\star\|.$$

As a result, applying inertial will slows down the performance.

Elliptical spiral trajectory For elliptical spiral, we consider the following form of M

$$M = \cos(\theta)\begin{bmatrix}\cos(\theta) & \frac{l}{s}\sin(\theta)\\ -\frac{s}{l}\sin(\theta) & \cos(\theta)\end{bmatrix},$$

for some $\theta \in]0, \pi/2[$ and $l, s > 0$. The property of ϑ_k becomes more complicated for the elliptical spiral, as ϑ_k varies in an interval $[\underline{\vartheta}, \overline{\vartheta}] \subset]0, \pi[$. Though the expressions of $\underline{\vartheta}, \overline{\vartheta}$ can be obtained explicitly based on the result of Section A.3, here we only provide descriptive explanation.

As we can observe from Figure 7 (d), that the angle ϑ_k varies in an interval $[\underline{\vartheta}, \overline{\vartheta}]$ where $\underline{\vartheta} < \pi/2$ and $\overline{\vartheta} > \pi/2$. This means that the direction $z_k - z_{k-1}$ only points towards $z^\star$ for *acute* ϑ_k. In turn, inertial provides acceleration when k are such that ϑ_k is acute and does not for the others. As a result, the overall performance of inertial is not clear in general.

Remark 4.2. It should be noted that the above discussion is in $\mathbb{R}^2$, the conclusion obtained for the spiral trajectories cannot be directly extended to higher dimension. For example, for logarithmic spiral, instead of being equal to $\pi/2$ for all k, ϑ_k is a varying sequence that converges to $\pi/2$. But still, as the trajectory of the sequence eventually settles on $\mathbb{R}^2$ (if the leading eigenvalue is unique), the above discussions hold true.

5 A^{2}FoM: adaptive acceleration for first-order methods

The trajectory property implies that, the sequence generated by first-order method eventually settles onto a regular path, *i.e.* straight line or spiral. In turn, we can use such regularity to design adaptive acceleration for first-order methods, which is called "A^{2}FoM" and described in Algorithm 5.

5.1 Trajectory following adaptive acceleration

We describe how to use the regularity of the trajectory to design a linear prediction scheme for acceleration. Recall the general inertial scheme (1.2) for first-order method

$$\bar{z}_k = \mathscr{E}(\bar{z}_{k-1}, z_k, z_{k-1}, ...),$$
$$z_{k+1} = \mathscr{F}(\bar{z}_k, z_k, z_{k-1}, ...).$$

From our discussion in the last section, to provide acceleration, the extrapolation operator $\mathscr{E}$ should be able to adapt itself to the trajectory of the sequence $\{z_k\}_{k\in\mathbb{N}}$. To this end, we propose a *trajectory following linear prediction strategy*, which locally fits the trajectory of the sequence $\{z_k\}_{k\in\mathbb{N}}$ and predict the future points. The basic idea of linear prediction is: let $q \in \mathbb{N}_+$ be a positive integer, given $\{z_{k-j}\}_{j=0}^{q+1}$ and $v_j \stackrel{\text{def}}{=} z_j - z_{j-1}$, forecast the future iterates by considering how the past directions $v_{k-1}, \dots, v_{k-q}$ approximate the latest direction v_k. More precisely,

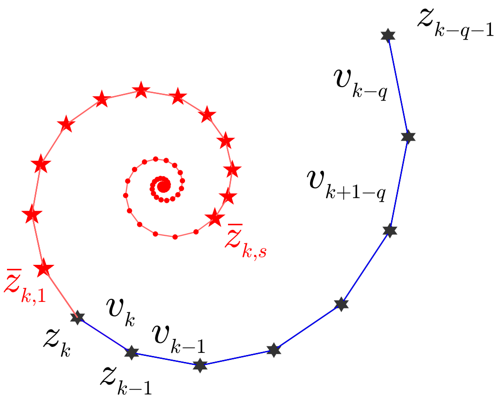


Figure 8: Illustration of linear prediction.

- First use $\{v_{k-j}\}_{j=1}^{q}$ to represent v_k, which is a least square problem. Denote $V_{k-1} \stackrel{\text{def}}{=} [v_{k-1}, \cdots, v_{k-q}] \in \mathbb{R}^{n\times q}$, and let

  $$c_k \in \mathrm{Argmin}_{c\in\mathbb{R}^q} \|V_{k-1}c - v_k\|^2 = \|\textstyle\sum_{j=1}^{q} c_j v_{k-j} - v_k\|^2.$$

  Then we have $v_k \approx V_{k-1}c_k$, and the representation if perfect if $v_k \in \mathrm{ran}(V_{k-1})$.
- Suppose we know z_{k+1}, then follow the first step we have $v_{k+1} \approx V_k c_{k+1}$. Since the trajectory locally is regular, we have $c_{k+1} \approx c_k$, this means we have the approximation $v_{k+1} \approx V_k c_k$. As a result, we obtain an approximation of z_{k+1} which is

  $$z_{k+1} \approx \bar{z}_{k,1} \stackrel{\text{def}}{=} z_k + V_k c_k.$$
- By iterating the second step s times where s is any positive integer, we obtain an approximation of z_{k+s} which is $\bar{z}_{k,s} \approx z_{k+s}$.

In Figure 8, we provide a graphical illustration of linear prediction: black dots are the given $q+2$ points, and red star points are the outputs of linear prediction from 1-step prediction to s-step prediction. If we run the prediction until $s=+\infty$, we obtain a complete spiral.

It can be observed that the above procedure is totally linear, therefore we can derive a simple matrix representation for linear prediction. Given a vector $c \in \mathbb{R}^q$, define the mapping H by

$$H(c) = \left[\begin{array}{c|c} c_{1:q-1} & \mathrm{Id}_{q-1} \\ c_q & 0_{1,q-1} \end{array} \right] \in \mathbb{R}^{q \times q}. \tag{5.1}$$

Let $C_k = H(c_k)$, note that $V_k = V_{k-1}C_k$. Denote $\bar{V}_{k,0} \overset{\text{def}}{=} V_k$ and for $s \geq 1$, define

$$\bar{V}_{k,s} \overset{\text{def}}{=} \bar{V}_{k,s-1}C_k \overset{\text{def}}{=} V_k C_k^s,$$

where C_k^s is the power of C_k. Let $(C)_{(:,1)}$ be the first column of matrix C, then

$$\bar{z}_{k,s} = z_k + \textstyle\sum_{i=1}^s (\bar{V}_{k,i})_{(:,1)} = z_k + \sum_{i=1}^s V_k (C_k^i)_{(:,1)} = z_k + V_k \big(\sum_{i=1}^s C_k^i\big)_{(:,1)}, \tag{5.2}$$

which is the desired trajectory following extrapolation scheme. Now define the extrapolation parameterized by s,q as

$$\mathscr{E}_{s,q}(z_k, \cdots, z_{k-q-1}) \overset{\text{def}}{=} V_k \big(\textstyle\sum_{i=1}^s C_k^i\big)_{(:,1)},$$

we obtain the following trajectory following adaptive acceleration for first-order method.

```
Algorithm 5: A^2FoM: Adaptive Acceleration for First-order Methods
    Input: Let $s \geq 1, q \geq 1$ be integers.
    Initial: Let $\bar{z}_0 = z_0 \in \mathbb{R}^n$ and $V_0 = 0 \in \mathbb{R}^{n \times q}$.
    Repeat:
        • If $\mathrm{mod}(k, q+2) = 0$: Compute $C_k$ as described above, if $\rho(C_k) < 1$:
              $\bar{z}_k = z_k + a_k \mathscr{E}_{s,q}(z_k, \cdots, z_{k-q-1})$.
        • If $\mathrm{mod}(k, q+2) \neq 0$: $\bar{z}_k = z_k$.
        • For $k \geq 1$:
              $z_{k+1} = \mathscr{F}(\bar{z}_k),\ v_{k+1} = z_{k+1} - z_k$  and  $V_{k+1} = [v_{k+1}|v_k|\cdots|v_{k-q+2}]$.
    Until: $\|v_k\| \leq \mathrm{tol}$.
```

Remark 5.1.

- When $\mathrm{mod}(k, q+2) \neq 0$, one can also consider $\bar{z}_k = z_k + a_k(z_k - z_{k-1})$ with properly chosen a_k. In stead of every $q+2$ steps, one can also consider $q+i$ with $i \geq 2$.
- A^{2}FoM carries out $q+2$ standard FoM iterations to set up the extrapolation step $\mathscr{E}_{s,q}$. As $\mathscr{E}_{s,q}$ contains the sum of the powers of C_k, it is guaranteed to be convergent if $\rho(C_k) < 1$. Therefore, we only apply $\mathscr{E}_{s,q}$ when the spectral radius $\rho(C_k) < 1$ is true. In this case, there is a closed form expression for $\mathscr{E}_{s,q}$ when $s = +\infty$; See Eq. (6.3).
- The purpose of adding a_k in front of $\mathscr{E}_{s,q}(z_k, \cdots, z_{k-q-1})$ is so that we can control the value of a_k to ensure the convergence of the algorithm; See Section 5.2.
- Though in this paper, we restricted ourselves with finite dimensional Euclidean space, our Algorithm 5 and its global convergence (Theorem 5.3) are readily extended to general Hilbert space.

Remark 5.2. In (5.2) we need to consider the sum of the power of C_k,

$$S_s = \textstyle\sum_{i=1}^s C_k^i.$$

Suppose that $\mathrm{Id}-C_k$ is invertible, recall the Neumann series $(\mathrm{Id}-C_k)^{-1}=\sum_{i=0}^{+\infty}C_k^i$. Therefore, for $s=+\infty$,

$$S_{+\infty}=(\mathrm{Id}-C_k)^{-1}-\mathrm{Id}=C_k(\mathrm{Id}-C_k)^{-1}. \tag{5.3}$$

In turn, for the finite s, we have $S_s=(C_k-C_k^{s+1})(\mathrm{Id}-C_k)^{-1}$.

5.2 Convergence of A^{2}FoM

In this part we study the global convergence property of A^{2}FoM . We first show that the A^{2}FoM can be treated as a perturbation of the original fixed-point iteration, and then discuss its convergence properties. Let $\varepsilon_k\in\mathbb{R}^n$ whose value takes

$$\varepsilon_k=\begin{cases} 0:\mathrm{mod}(k,q+2)\neq 0 \ \text{ or } \ \mathrm{mod}(k,q+2)=0 \ \& \ \rho(C_k)\geq 1,\\ a_k\mathscr{E}_{s,q}(z_k,\cdots,z_{k-q-1}):\mathrm{mod}(k,q+2)=0 \ \ \& \ \ \rho(C_k)<1. \end{cases}$$

Then the Algorithm 5 can be written as

$$z_{k+1}=\mathscr{F}(z_k+\varepsilon_k). \tag{5.4}$$

Based on the above reformulation, we have the following convergence result for Algorithm 5 which is based on the classic convergence result of inexact Krasnosel'skiĭ-Mann fixed-point iteration [7, Proposition 5.34].

Theorem 5.3. *For Algorithm 5, suppose that the fixed-point operator $\mathscr{F}$ is averaged non-expansive whose set of fixed-points is non-empty. If the perturbation error is absolutely summable, i.e. $\sum_k\|\varepsilon_k\|<+\infty$, then there exists a $z^\star\in\mathrm{fix}(\mathscr{F})$ such that $z_k\to z^\star$.*

Proof. From (5.4), we have that

$$z_{k+1}=\mathscr{F}(z_k+\varepsilon_k)=\mathscr{F}(z_k)+\big(\mathscr{F}(z_k+\varepsilon_k)-\mathscr{F}(z_k)\big).$$

Given any $z^\star\in\mathrm{fix}(\mathscr{F})$, there holds

$$\|z_{k+1}-z^\star\|\leq\|\mathscr{F}(z_k)-\mathscr{F}(z^\star)\|+\|\mathscr{F}(z_k+\varepsilon_k)-\mathscr{F}(z_k)\|\leq\|z_k-z^\star\|+\|\varepsilon_k\|,$$

which means that $\{z_k\}_{k\in\mathbb{N}}$ is quasi-Fejér monotone with respect to $\mathrm{fix}(\mathscr{F})$. Then invoke [7, Proposition 5.34] we obtain the convergence of the sequence $\{z_k\}_{k\in\mathbb{N}}$. □

Remark 5.4. The perturbation perspective of A^{2}FoM implies that we can incorporate other errors in the iteration, as long as the error is absolutely summable. Such as $\mathscr{F}(\bar{z}_k)$ is computed approximately, and the accuracy is increasing along the iteration.

The above convergence result indicates that we need a proper strategy to ensure that $a_k\mathscr{E}_{s,q}(z_k,\cdots,z_{k-q-1})$ is absolutely summable. This can be obtained via the following safeguard strategy, which is inspired by [4].

```
Algorithm 6: A$^2$FoM with safeguard
  Input: Let $a,b,\delta>0$ and $s\geq 1, q\geq 1$ be integers.
  Initial: Let $z_0\in\mathbb{R}^n$ and $\bar{z}_0=z_0$, set $V_0=0\in\mathbb{R}^{n\times q}$;
  Repeat:
    • If $\mathrm{mod}(k,q+2)=0$: Compute $C_k$ as described above, if $\rho(C_k)<1$:
        $a_k=\min\Big\{a,\frac{b}{k^{1+\delta}\|\mathscr{E}_{s,q}(z_k,\cdots,z_{k-q-1})\|}\Big\}$,
        $\bar{z}_k=z_k+a_k\mathscr{E}_{s,q}(z_k,\cdots,z_{k-q-1})$.
    • If $\mathrm{mod}(k,q+2)\neq 0$:  $\bar{z}_k=z_k$.
    • For $k\geq 1$:
        $z_{k+1}=\mathscr{F}(\bar{z}_k)$,  $v_{k+1}=z_{k+1}-z_k$  and  $V_{k+1}=[v_{k+1}|v_k|\cdots|v_{k-q+2}]$.
  Until: $\|v_k\|\leq\mathrm{tol}$.
```

6 Acceleration guarantees of A^{2}FoM

As our A^{2}FoM is motivated by the local trajectory of the sequence $\{z_k\}_{k\in\mathbb{N}}$, in this part we turn to the local perspective and study the local acceleration guarantees of A^{2}FoM . We first recall several well established vector extrapolation methods in the field of numerical analysis, build connection with our linear prediction and then discuss the acceleration guarantees.

6.1 Vector extrapolation techniques

Vector extrapolation techniques provide a generic recipe for the acceleration of sequences, without specific knowledge of how the sequence is generated.

In the following, we describe two popular techniques for vector extrapolation of a sequence $\{x_k\}_k$. Let $u_k \overset{\text{def}}{=} x_{k+1} - x_k$ and define the matrix

$$U_j \overset{\text{def}}{=} \left[u_k | u_{k+1} | \cdots | u_{k+j}\right]. \tag{6.1}$$

Idea of vector extrapolation methods Suppose we are observing a sequence $\{x_k\}_k$ generated by

$$x_{k+1} = T x_k + d \tag{6.2}$$

where T is a matrix and d is a vector, which are possibly unknown. Assume that $\rho(T) < 1$ so that $\lim_{k\to+\infty} x_k \overset{\text{def}}{=} x^\star$ exists. We say that P is a minimal polynomial of T with respect to a vector v if it is the monic polynomial of least degree such that $P(T)v = 0$.

It is known [75] that if $P(\lambda) = \sum_{i=0}^r c_i \lambda^i$ is a minimal polynomial of T with respect to $x_k - x^\star$, it is also minimal with respect to $x_{k+1} - x_k$. Moreover, $\sum c_i \neq 0$ and $x^\star = \frac{1}{\sum_i c_i} \sum_{i=0}^r c_i u^{k+i}$ where $u_k \overset{\text{def}}{=} x_{k+1} - x_k$. Therefore, we can compute $x^\star$ from finitely many values of this sequence provided that the minimal polynomial coefficients are known. To compute these coefficients, note that

$$0 = P(T) u_k = \sum_{i=0}^r c_i T^i u_k = \sum_{i=0}^r c_i u^{k+i}.$$

Since $c_r = 1$, we can write this equation as $U_r c = 0$ and $U_{r-1} c' = -u^{k+r}$, where $c' = (c_0, \ldots c_{r-1})^\top$. Note that this is an overdetermined system if $r \leq d$, and is consistent and has a unique solution. Finally, setting $\gamma_i = \frac{c_i}{\sum c_i}$, we have $x^\star = \sum_i \gamma_i x_{k+i}$. This process of computing the coefficients is known as minimal polynomial extrapolation, and is summarized below.

Algorithm 7: Minimal polynomial extrapolation (MPE)

```
1. Choose integers $r$ and $k$ and input the vectors $x_k, x_{k+1}, \ldots, x_{k+r+1}$.
2. Compute the vectors $u_k, u_{k+1}, \ldots, u_{k+r}$ and the matrix $U_{r-1} = \left[u_k | u_{k+1} | \cdots | u_{k+r-1}\right]$.
3. Find $c' \overset{\text{def}}{=} \left[c_0, \ldots, c_{r-1}\right]^\top$ as the least squares solution to $U_{r-1} c' = -u^{k+r}$. Set $c_r \overset{\text{def}}{=} 1$ and
   $\gamma_i = c_i / \sum_{j=0}^r c_j$ for $i = 0, \ldots, r$.
4. Compute $s_{k,r} \overset{\text{def}}{=} \sum_{i=0}^r \gamma_i x_{k+i}$ as an approximation to $\lim_{k\to+\infty} x_k = x^\star$.
```

In general, if r is chosen too small, $\sum_i c_i$ might be zero and MPE will fail. To circumvent this, reduced rank extrapolation was introduced, where step 3 is replaced with a constrained minimization problem.

Algorithm 8: Reduced rank extrapolation (RRE)

```
1. Choose integers $r$ and $k$ and input the vectors $x_k, x_{k+1}, \ldots, x_{k+r+1}$.
2. Compute the vectors $u_k, u_{k+1}, \ldots, u_{k+r}$ and for the matrix $U_r$.
3. Let $\gamma \in \mathrm{argmin}_\gamma \|U_r \gamma\|$ subject to $\sum_{i=0}^r \gamma_i = 1$.
4. Compute $s_{k,r} \overset{\text{def}}{=} \sum_{i=0}^r \gamma_i x_{k+i}$ as an approximation to $\lim_{k\to+\infty} x_k = x^\star$.
```

Another form of convergence acceleration technique is Anderson acceleration [5], whose formulation is similar to that of RRE (and is equivalent in the linear setting), further details about its relation to vector extrapolation technique can be found in [18]. There has been recent work on applying this extrapolation technique to accelerate first order algorithms [71, 80, 63]. One of the challenges of applying these methods is that while $s_{k,r} \to s$ as $r \to \infty$, choosing large values for r could lead to U_r being ill-conditioned, [71] suggested to circumvent this issue using regularization techniques when solving step 3. However, naive regularization could actually slow down convergence, and an adaptive choice of the regularization parameter may lead to many evaluations of the objective function which may be costly.

6.2 Equivalence between A^{2}FoM and MPE

We build connection between our A^{2}FoM with MPE/RRE for the case of $s = +\infty$, that is when the linear prediction is taken for infinite steps.

Owing to (5.3), when $s = +\infty$, from (5.2) we get

$$\begin{aligned}\bar{z}_{k,\infty} \overset{\text{def}}{=} z_k + V_k\big((\mathrm{Id} - C_k)^{-1} - \mathrm{Id}\big)_{(:,1)} &= z_k - v_k + V_k\big((\mathrm{Id} - C_k)^{-1}\big)_{(:,1)} \\ &= z_{k-1} + V_k\big((\mathrm{Id} - C_k)^{-1}\big)_{(:,1)} \\ &= \tfrac{1}{1-\sum_{i=1}^{s} c_{k,i}}\big(z_k - \textstyle\sum_{j=1}^{q-1} c_{k,j} z_{k-j}\big),\end{aligned} \tag{6.3}$$

which turns out to be MPE, with the slight difference of taking the weighted sum of $\{z_j\}_{j=k-q+1}^{k}$ as opposed to the weighted sum of $\{z_j\}_{j=k-q}^{k-1}$. Note that if the coefficients c is computed in the following way: $b \in \operatorname{argmin}_{b\in\mathbb{R}^{q+1}, \sum_j b_j = 1} \|\sum_{j=0}^{q} b_j v_{k-j}\|$ and $b_0 \neq 0$ and define $c_j \overset{\text{def}}{=} -b_j/b_0$ for $j = 1, \ldots, q$. Then,

$$\big(1 - \textstyle\sum_{i=1}^{q} c_i\big)^{-1} = \frac{b_0}{b_0 + \sum_{j=1}^{q} b_j} = b_0,$$

and $\bar{z}_{k,\infty} = \sum_{j=0}^{q-1} b_j z_{k-j}$ is precisely the RRE update (again with the slight difference of summing over iterates shifted by one iteration).

Remark 6.1. Based on the structure of C in (5.1), simple calculation yields

$$\mathrm{Id} - C = \begin{bmatrix} (1-c_1) & -1 & 0 & \cdots & 0 \\ -c_2 & 1 & -1 & \ddots & \vdots \\ -c_3 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & -1 \\ -c_q & 0 & \cdots & 0 & 1 \end{bmatrix}_{q\times q} \quad \text{and} \quad (\mathrm{Id} - C)^{-1} = \frac{1}{1-\sum_{i=1}^{s} c_i} \begin{bmatrix} 1 & 1 & 1 & \cdots & \cdots & 1 & 1 \\ b_2 & \bar{b}_2 & \bar{b}_2 & \cdots & \cdots & \bar{b}_2 & \bar{b}_2 \\ b_3 & b_3 & \bar{b}_3 & \cdots & \cdots & \bar{b}_3 & \bar{b}_3 \\ b_4 & b_4 & b_4 & \bar{b}_4 & \cdots & \bar{b}_4 & \bar{b}_4 \\ \vdots & & & & & & \vdots \\ b_q & b_q & \cdots & \cdots & \cdots & b_q & \bar{b}_q \end{bmatrix}_{q\times q}$$

where $b_j \overset{\text{def}}{=} \sum_{i\geq j} c_i$ and $\bar{b}_j = 1 - \sum_{i<j} c_i$ such that $\bar{b}_j - b_j = 1 - \sum_j c_j$.

6.3 Acceleration guarantees of A^{2}FoM

We are now ready to discuss the acceleration guarantees of A^{2}FoM . We first characterize the prediction error of our proposed A^{2}FoM , and then discuss its acceleration guarantees based on the relation with MPE/RRE.

6.3.1 Prediction error of A^{2}FoM

To discuss the prediction error of A^{2}FoM , we need to rewrite (3.1) first. Denote $f_k = o(\|z_k - z_{k-1}\|)$ and

$$F_k = [f_k | f_{k-1} | \cdots | f_{k-q+1}] \in \mathbb{R}^{n\times q}.$$

Recall that $v_k \overset{\text{def}}{=} z_k - z_{k-1}$ and $V_k = [v_k | v_{k-1} | \cdots | v_{k-q+1}] \in \mathbb{R}^{n\times q}$, from (3.1) we have $v_k = M_{\mathscr{F}}(v_{k-1}) + f_{k-1}$ and

$$V_k = M_{\mathscr{F}} V_{k-1} + F_{k-1}. \tag{6.4}$$

By virtue the definition of the coefficients matrix $C_k \overset{\text{def}}{=} H(c_k)$ of (5.1), define $E_{k,j} \overset{\text{def}}{=} V_k C_k^j - V_{k+j}$ for $j \geq 1$ and

$$E_{k,0} \overset{\text{def}}{=} V_{k-1}C_k - V_k = \begin{bmatrix}(V_{k-1}c_k - v_k) & 0 & \cdots & 0\end{bmatrix}. \tag{6.5}$$

We arrive at the following relation between the extrapolated point $\bar{z}_{k,s}$ and the $(k+s)$'th point of $\{z_k\}_{k\in\mathbb{N}}$

$$\bar{z}_{k,s} = z_k + \textstyle\sum_{j=1}^{s}(v_{j+k} + (E_{k,j})_{(:,1)}) = z_{k+s} + \sum_{j=1}^{s}(E_{k,j})_{(:,1)}.$$

As a result, we derive the following proposition on the prediction error $\bar{z}_{k,s} - z^\star$.

Theorem 6.2 (Prediction error). *Given a first-order method of the form (1.1), let (3.1) be its local linearization. For Algorithm 5, when the linear prediction is applied, we have the following error bounds: Let*

$$B_{k,s} \overset{\text{def}}{=} \max_{i\in\{0,1\}, t\leq s} \{\|\textstyle\sum_{\ell=i}^{t} M_{\mathscr{F}}^{\ell}\|, \|\sum_{\ell=i}^{t} C_k^{\ell}\|\}$$

and define the coefficients fitting error as $\varepsilon_k \overset{\text{def}}{=} \|\sum_{i=1}^{q} c_{k,i}v_{k-i} - v_k\|$. Then, the prediction error $\bar{z}_{k,s} - z^\star$ satisfies

$$\|\bar{z}_{k,s} - z^\star\| \leq \|M_{\mathscr{F}}^{s}(z_k - z^\star)\| + \|\textstyle\sum_{\ell=0}^{s-1} M_{\mathscr{F}}^{\ell}\|\|\hat{f}_k\| + B_{k,s}(\varepsilon_k + \|F_{k-1}\|),$$

where $\hat{f}_k \overset{\text{def}}{=} M_{\mathscr{F}}(z_{k-1} - z^\star) - (z_k - z^\star)$.

In the case of $s = +\infty$, there holds

$$\|\bar{z}_{k,+\infty} - z^\star\| \leq \|(\mathrm{Id} - M_{\mathscr{F}})^{-1}\|\|\hat{f}_k\| + \varepsilon_k \frac{\sum_{\ell=1}^{+\infty}\|M_{\mathscr{F}}^{\ell}\|}{1 - \sum_i c_{k,i}} + \sum_{\ell=0}^{+\infty}\|M_{\mathscr{F}}^{\ell}\|\|F_{k-1}((\mathrm{Id} - C_k)^{-1} - \mathrm{Id})_{(:,1)}\|.$$

The proof of the theorem can be found in Section C of the appendix.

Remark 6.3.

- The fact that $B_{k,s}$ is uniformly bounded in s if $\rho(M_{\mathscr{F}}) < 1$, and $\rho(C_k) < 1$ follows because this implies that $\sum_{\ell=1}^{+\infty}\|M_{\mathscr{F}}^{\ell}\| < +\infty$ thanks to the Gelfand formula, and $\sum_{i=0}^{+\infty} C_k^i = (\mathrm{Id} - C_k)^{-1}$ and its $(1,1)^{th}$ entry is precisely $\frac{1}{1-\sum_i c_{k,i}}$.
- In Theorem 6.2, the prediction error consists of two main sources: coefficient fitting error of $E_{k,0}$ and linearization error of F_{k-1} which corresponds to the small o-terms. When the small o-term in (3.1) vanishes, that is $F_{k-1} = 0$, then it follows from the proof that

  $$\|\bar{z}_{k,s} - z^\star\| \leq \|z_{k+s} - z^\star\| + B_{k,s}\varepsilon_k$$

  and if the spectral radius $\rho(M_{\mathscr{F}}) < 1$ and $\rho(C_k) < 1$, then

  $$\|\bar{z}_{k,+\infty} - z^\star\| \leq \textstyle\sum_{\ell}\|M_{\mathscr{F}}^{\ell}\|\frac{\varepsilon_k}{1-\sum_i c_{k,i}}.$$

6.3.2 Acceleration guarantees

As shown in Theorem 6.2, a key quantity governing the amount of acceleration is the coefficient fitting error ε_k. For the case that the small o-terms vanish, this error can be bounded using existing results of vector extrapolation. In the following, we assume that (1.1) can be linearized without small o-term and derive acceleration guarantees for Algorithm 5.

Theorem 6.4 (Acceleration guarantees). *Given a first-order method of the form (1.1), suppose there exists a linear matrix $M_{\mathscr{F}}$ such that it can be linearized of the form*

$$z_{k+1} - z_k = M_{\mathscr{F}}(z_k - z_{k-1}).$$

Suppose that $M_{\mathscr{F}}$ is diagonalizable. Let $\{\lambda_j\}_j$ denote its distinct eigenvalues ordered such that $|\lambda_j| \geq |\lambda_{j+1}|$ and $|\lambda_1| = \rho(M_{\mathscr{F}}) < 1$. Suppose that $|\lambda_q| > |\lambda_{q+1}|$. Then we have the following bounds on ε_k

- *Asymptotic bound (fixed q and as $k \to +\infty$): $\varepsilon_k = O(|\lambda_{q+1}|^k)$.*

- *Non-asymptotic bound (fixed q and k): Suppose that $\lambda(M_{\mathscr{F}})$ is real-valued and contained in the interval $[\alpha, \beta]$ with $-1 < \alpha < \beta < 1$. Then,*

$$\frac{\varepsilon_k}{1-\sum_i c_{k,i}} \leq K\beta^{k-q}\big(\tfrac{\sqrt{\eta}-1}{\sqrt{\eta}+1}\big)^q \tag{6.6}$$

  where $K \overset{\text{def}}{=} 2\|z_0 - z^\star\|\|(\mathrm{Id}-M)^{\frac{1}{2}}\|$ and $\eta = \frac{1-\alpha}{1-\beta}$.

Remark 6.5.

- As we have seen in Section 3, when R and J are both polyhedral, when the optimization problem is locally polyhedral around the solution, the small o-term vanishes and we have a perfect local linearization. Hence, the conditions of Theorem 6.2 holds for all k large enough.
- The first bound (i) shows that the extrapolated point $\bar{z}_{k,s}$ moves along the true trajectory as s increases, up to the fitting error ε_k. Although $\bar{z}_{k,+\infty}$ is essentially an MPE update which is known to satisfy error bound (6.6); See for instance [75]. This theorem offers a further geometric interpretation of these extrapolation methods in terms of following the "sequence trajectory", and combined with our local analysis of **FoM**, provides justification of these methods for the acceleration of non-smooth optimization problems.

Remark 6.6 (Acceleration guarantee and the choice of q).

- For Forward–Backward splitting method, as the angle θ_k converges to 0. For the coefficient fitting error we have $\varepsilon_k = o(\rho(M_{\mathscr{F}}))$, which indicate that A^{2}FoM can provide acceleration with $q = 1$. Since $q = 1$ corresponds to the inertial scheme, our result complies with the current literature on inertial Forward–Backward splitting methods.
- Theorem 6.4 (ii) shows that extrapolation improves the convergence rate from $O(|\lambda_1|^k)$ to $O(|\lambda_{q+1}|^k)$, and the non-asymptotic bound shows that the improvement of extrapolation is optimal in the sense of Nesterov [59]. Take Douglas–Rachford splitting for example, in the case of two non-smooth polyhedral terms, we must have $|\lambda_{2j-1}| = |\lambda_{2j}| > |\lambda_{2j+1}|$ for all $j \geq 1$. Hence, no acceleration can be guaranteed or observed when $q = 1$, while the choice of $q = 2$ provides guaranteed acceleration.

Remark 6.7 (Dealing with small o-terms). We now consider the coefficients fitting error of the perturbed problem $v_k = M_{\mathscr{F}}(v_{k-1}) + f_{k-1}$. Let $v_k^0 = M_{\mathscr{F}}(v_{k-1}^0)$ with $v_{k-q}^0 = v_{k-q}$ and let $c_k^0 \in \mathbb{R}^q$ and $C_k^0 = H(c_k^0)$ be the associated coefficients and coefficients matrix. Let ε^0 be the coefficients fitting error for this unperturbed problem, then

$$\begin{aligned}
\varepsilon = \min_{c\in\mathbb{R}^q} \|\textstyle\sum_{j=1}^q c_j v_{k-j} - v_k\| &\leq \varepsilon^0 + \|\textstyle\sum_{j=1}^q c_{k,j}^0 (v_{k-j} - v_{k-j}^0) - v_k - v_k^0\| \\
&\leq \varepsilon^0 + \|\textstyle\sum_{i=1}^q c_{k,i}^0 \big(\sum_{\ell=1}^{q-i} M_{\mathscr{F}}^{\ell-1} f_{k-i-\ell}\big) - \big(\sum_{\ell=1}^q M_{\mathscr{F}}^{\ell-1} f_{k-\ell}\big)\| \\
&\leq \varepsilon^0 + (1 + \|c^0\|_1) \max_{i=1}^q \|f_{k-i}\| \textstyle\sum_{\ell=1}^q \|M_{\mathscr{F}}^\ell\|
\end{aligned}$$

where we have used

$$v_{k-i} - v_{k-i}^0 = M_{\mathscr{F}}^{q-i}(v_{k-q} - v_{k-q}^0) + \sum_{\ell=1}^{q-i} M_{\mathscr{F}}^{\ell-1} f_{k-i-\ell} = \sum_{\ell=1}^{q-i} M_{\mathscr{F}}^{\ell-1} f_{k-i-\ell}.$$

Therefore, even with the presence of small o-terms, the coefficients fitting error can be bounded in terms of the small o-terms and the coefficients fitting error under exact linearization.

7 Implementation and numerical experiments

Our proposed adaptive acceleration scheme Algorithm 5 is quite abstract in the sense it is only presented for fixed-point iteration. While for first-order methods, as we have seen in Section 3, each method has a unique fixed-point characterization. Therefore, in section we discuss how to implement A^{2}FoM for different algorithms and provide numerical tests to demonstrate the performance of our acceleration scheme.

7.1 Forward–Backward splitting

We start with the Forward–Backward splitting algorithm, adapt A^{2}FoM to this method we obtain the following adaptive accelerated Forward–Backward splitting scheme.

```
Algorithm 9: A$^2$FB : Adaptive Acceleration for Forward–Backward splitting
  Input: $\gamma \in ]0, 2/L[$. Let $s \geq 1, q \geq 1$ be integers. Let $\theta > 0$.
  Initial:$\bar{x}_0 = x_0 \in \mathbb{R}^n$. Set $V_0 = 0 \in \mathbb{R}^{n \times q}$.
  Repeat:
    • If $\mathrm{mod}(k, q+2) = 0$: Compute $C_k$ via (5.1), if $\rho(C_k) < 1$ and $\angle(v_k, \mathscr{E}_{s,q}(x_k, \cdots, x_{k-q-1})) \leq \theta$:
          $\bar{x}_k = x_k + a_k \mathscr{E}_{s,q}(x_k, \cdots, x_{k-q-1})$.
    • If $\mathrm{mod}(k, q+2) \neq 0$: $\bar{x}_k = x_k$.
    • For $k \geq 1$:
          $x_{k+1} = \mathrm{prox}_{\gamma R}\big(\bar{x}_k - \gamma \nabla F(\bar{x}_k)\big)$,
          $v_{k+1} = x_{k+1} - x_k$ and $V_{k+1} = [v_{k+1} | v_k | \cdots | v_{k-q+2}]$.
  Until: $\|v_k\| \leq \mathrm{tol}$.
```

Remark 7.1. Note that for the above scheme, we have an extra check on the angle between $v_k = x_k - x_{k-1}$ and the extrapolated direction $\mathscr{E}_{s,q}(x_k, \cdots, x_{k-q-1})$, and the value of θ is close to 0. This is due to the fact that the trajectory of $\{x_k\}_{k \in \mathbb{N}}$ eventually is straight-line, we only accept $\mathscr{E}_{s,q}(x_k, \cdots, x_{k-q-1})$ if the angle $\angle(x_k, \mathscr{E}_{s,q}(x_k, \cdots, x_{k-q-1}))$ is small enough.

7.1.1 Least square problem

When $R = 0$, Forward–Backward splitting method recovers the gradient descent. For least square problem, gradient descent results in the linear system of (6.2). Therefore, in this example we compare the performance of the following methods:

- Gradient descent (GD), restarting FISTA [60].
- Our proposed scheme (LP) with $(q, s) = (6, +\infty)$.
- MPE, RRE and regularized non-linear acceleration (RNA) [71].

For each of the four extrapolation methods, we compare the basic scheme without line search (denoted as "basic") and adaptive one with line search (denoted as "LS"). For "RNA adaptive", not only line search, but also grid search on the regularization parameter is applied.

The following least square problem is considered

$$\min_{x \in \mathbb{R}^{50}} \frac{1}{2} \|Ax - f\|^2,$$

where three different choices of A are considered

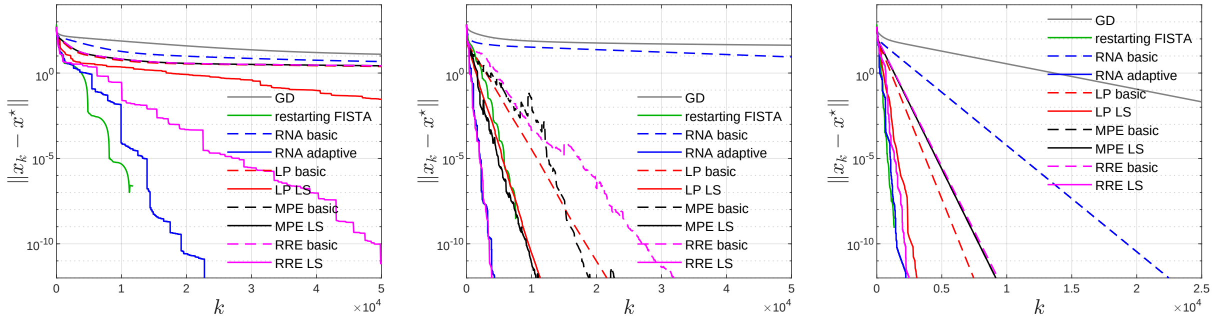
- Tridiagonal matrix with main diagonal elements equal to 2, and the elements of the first diagonal below and above main diagonal equal to -1;
- $A = \mathtt{rand}(51, 50)$ is generated from uniform distribution in $[-1, 1]$.
- $A = \mathtt{randn}(51, 50)$ is generated from normal Gaussian distribution.

The performance comparison of different methods are shown in Figure 9, from which we observe that

- Among all the algorithms, gradient descent (gray line) is the slowest, with "RNA basic" the second slowest.
- "RNA adaptive" shows the best overall performance, except for the tridiagonal A as it is slightly slower than "restarting FISTA". "restarting FISTA" is one of the fastest for tridiagonal and `randn` A, but not for `rand` A. "RRE LS" also performs very fast, particularly for `rand` and `randn` A.

- Our proposed scheme with infinity number prediction is faster than MPE but slower than RRE.

We remark that the performance of "restarting FISTA" is rather impressive given its simplicity and easy implementation.



(a) A is a tridiagonal matrix (b) $A = \texttt{rand}(50,50)$ (c) A = $\texttt{randn}$(50, 50)

Figure 9: Performance comparison among different schemes under different choices of A.

7.1.2 LASSO-type problem

Next we consider regularized least square problem of the form

$$\min_{x\in\mathbb{R}^n} R(x) + \frac{1}{2}\|\mathcal{K}x - f\|^2, \tag{7.1}$$

where R is regularization term, and $\mathcal{K} \in \mathbb{R}^{m\times n}$ is drawn from random Gaussian ensemble. f is the observation of some $\mathring{x}$ under $\mathcal{K}$ contaminated by noise w,

$$f = \mathcal{K}\mathring{x} + w. \tag{7.2}$$

In this experiment, three different cases of R are considered: sparsity promoting ℓ_1-norm, group sparsity promoting $\ell_{1,2}$-norm and low-rank promoting nuclear norm. The detailed settings of each example are

ℓ_1-norm $(m,n) = (48,128)$, $\mathring{x}$ has 8 non-zero elements.
$\ell_{1,2}$-norm $(m,n) = (48,128)$, $\mathring{x}$ has 3 non-zero blocks of size 4.
Nuclear norm $(m,n) = (640,1024)$, $\mathring{x} \in \mathbb{R}^{32\times 32}$ and $\mathrm{rank}(\mathring{x}) = 4$.

The following scheme are compared

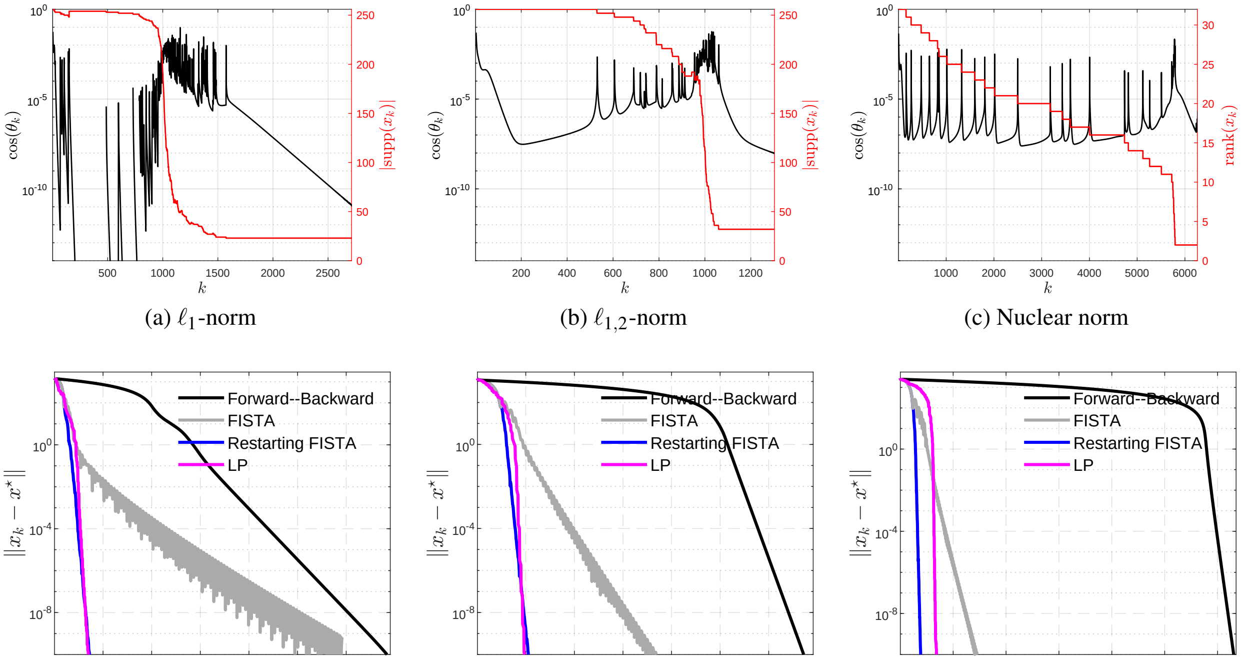
- Forward–Backward splitting, FISTA and restarting FISTA.
- Our proposed scheme (LP) with $(q,s) = (4,+\infty)$.

The finite activity identification and $\cos(\theta_k)$ of Forward–Backward splitting is provided in the first row of Figure 10, the observations are quite close to those of Example 3.1. The performance comparison of the above methods is presented in the second row of Figure 10, and we observe that

- Similar to the least square example, Forward–Backward splitting method is the slowest one. However, note that in terms of local linear convergence rate, FISTA is the slowest one — see the local slope of the gray and black line. The problem of Forward–Backward splitting method is that it needs much longer time to identified the underlying manifold.
- Restarting FISTA (blue line) is the fastest among all methods, our proposed linear prediction is as fast as restarting FISTA for the first two examples and slightly slower for the last example.

7.2 Douglas–Rachford splitting

Now we turn to the Douglas–Rachford splitting method, for which we obtain the following adaptive scheme. A bit difference between the following iteration and that of (3.3) is that we start the iteration with x_k.



(a) ℓ_1-norm (b) $\ell_{1,2}$-norm (c) Nuclear norm

(d) ℓ_1-norm (e) $\ell_{1,2}$-norm (f) Nuclear norm

Figure 10: Comparison between methods for solving regularized least square.

```
Algorithm 10: A^2DR : Adaptive Acceleration for Douglas–Rachford splitting
    Input: $\gamma > 0$. Let $s \geq 1, q \geq 1$ be integers.
    Initial: $\bar{z}_0 = z_0 \in \mathbb{R}^n$, $x_0 = \mathrm{prox}_{\gamma J}(\bar{z}_0)$. Let $V_0 = 0 \in \mathbb{R}^{n \times q}$.
    Repeat:
        • If $\mathrm{mod}(k, q+2) = 0$: Compute $C_k$ via (5.1), if $\rho(C_k) < 1$: $\bar{z}_k = z_k + a_k \mathscr{E}_{s,q}(z_k, \cdots, z_{k-q-1})$.
        • If $\mathrm{mod}(k, q+2) \neq 0$: $\bar{z}_k = z_k$.
        • For $k \geq 1$:
              $x_k = \mathrm{prox}_{\gamma J}(\bar{z}_k)$,
              $u_{k+1} = \mathrm{prox}_{\gamma R}(2x_k - \bar{z}_k)$,
              $z_{k+1} = \bar{z}_k + u_{k+1} - x_k$,
              $v_{k+1} = z_{k+1} - z_k$  and  $V_{k+1} = [v_{k+1} | v_k | \cdots | v_{k-q+2}]$.
    Until: $\|v_k\| \leq \mathrm{tol}$.
```

7.2.1 Basis pursuit type problems

Now suppose that there is no noise in the observation model (7.2), *i.e.* $w = 0$. Then instead of solving (7.1), the following equality constrained problem should be considered

$$\min_{x \in \mathbb{R}^n} R(x) \quad \text{subject to} \quad \mathscr{K} x = \mathscr{K} \mathring{x}.$$

Furthermore, the above constrained problem can be ban be formulated as

$$\min_{x \in \mathbb{R}^n} R(x) + J(x), \tag{7.3}$$

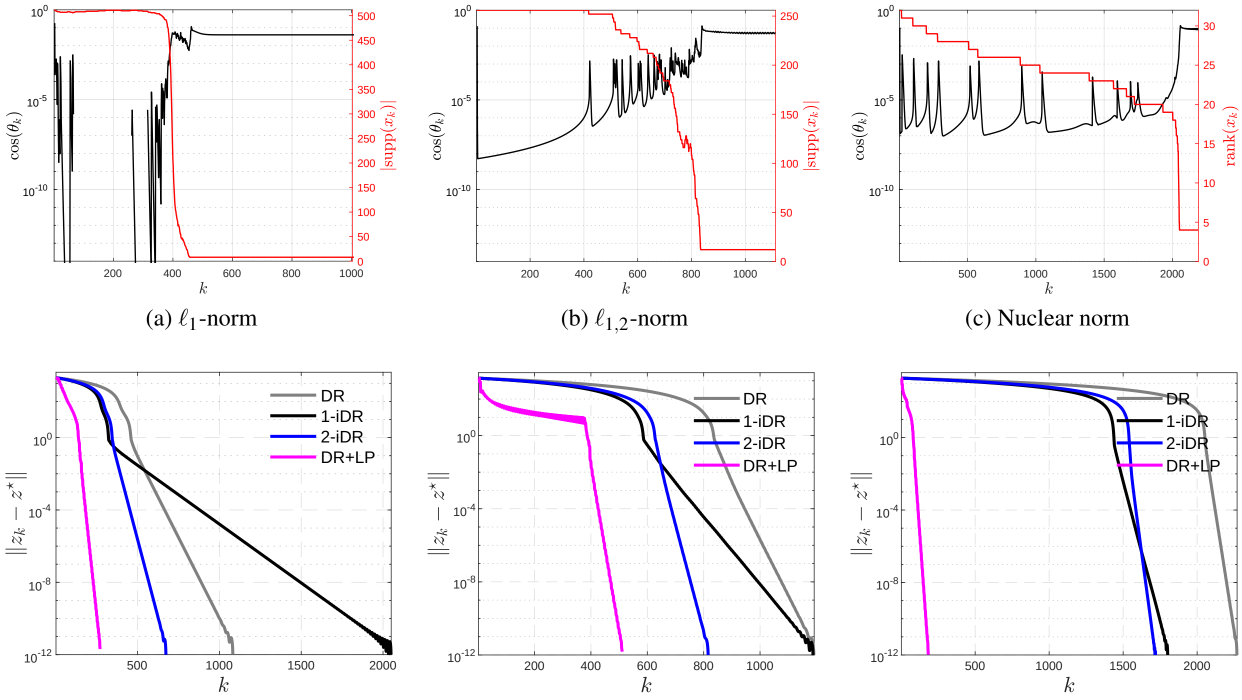
where $J = \iota_\Omega(\cdot)$ is the indicator function of the constraint $\Omega \overset{\text{def}}{=} \{x \in \mathbb{R}^n : \mathscr{K}\mathring{x} = \mathscr{K}x\} = \mathring{x} + \ker(\mathscr{K})$. As both functions R and J are non-smooth, a proper choice to solve (7.3) is the Douglas–Rachford splitting. The proximity operator of J is the projection operator onto Ω, which reads $\mathrm{prox}_{\gamma J}(x) = x + \mathscr{K}^+(f - \mathscr{K}x)$ where $\mathscr{K}^+ = \mathscr{K}^T(\mathscr{K}\mathscr{K}^T)^{-1}$ is the Moore-Penrose pseudo-inverse of $\mathscr{K}$. For R, again three examples are considered: $\ell_1, \ell_{1,2}$ and nuclear norm, and the setting of each example is the same as the Forward–Backward splitting experiments.

The following scheme are compared

- Douglas–Rachford splitting (DR), the standard two-point inertial DR (4.1) (1-iDR), the three-point inertial DR (4.3) (2-iDR).
- Our proposed scheme (LP) with $(q, s) = (4, 100), (4, +\infty)$.

The finite activity identification and $\cos(\theta_k)$ of Douglas–Rachford splitting is provided in the first row of Figure 11, the observations are quite close to those of Example 3.2. The performance comparison of the above methods is presented in the second row of Figure 11, and we observe that

- Only "2-iDR" shows constant better performance than DR. Locally, the convergence speed of "1-iDR" is the slowest among all schemes. This observation comply with our discussion in Section 4.3.
- Linear prediction is the fastest among all the schemes, especially for nuclear norm. The main improvement of LP is that it needs much shorter time to identify the manifolds.



(a) ℓ_1-norm

(b) $\ell_{1,2}$-norm

(c) Nuclear norm

(d) ℓ_1-norm

(e) $\ell_{1,2}$-norm

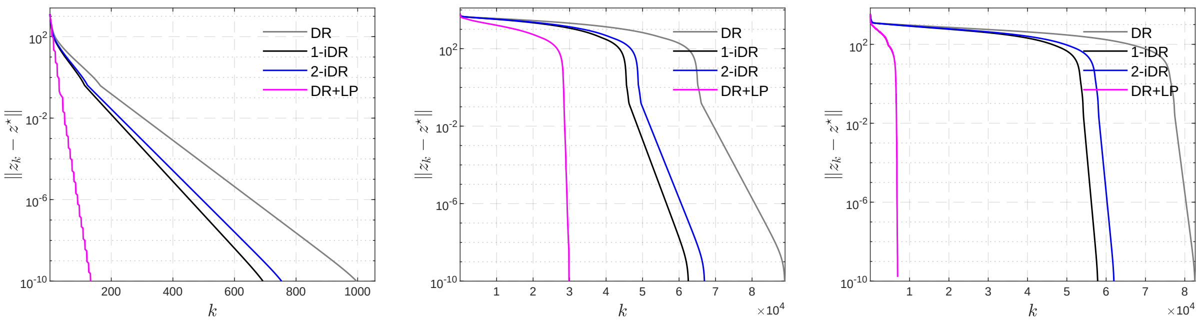
(f) Nuclear norm

Figure 11: Comparison between methods for solving basis pursuit type problem.

7.2.2 LASSO problem

We also consider the LASSO problem (4.4) to demonstrate the performance. Three data sets from LIBSVM[4] are considered: `australian`, `mushrooms` and `covtype`. The observation are shown in Figure 12, we can see that linear prediction shows clear advantages over the compared ones.

[4] https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/



(a) `australian` (b) `mushrooms` (c) `covtype`

Figure 12: Comparison of Douglas–Rachford schemes for solving LASSO problem.

7.3 Primal–Dual splitting

The third example we consider is the Primal–Dual splitting method. Adapt A^{2}FoM to the method, we obtain the following iteration. Note that the fixed-point sequence of Primal–Dual splitting method is the augmented variable z_k defined in (3.6).

```
Algorithm 11: A^2PD : Adaptive Acceleration for Primal–Dual splitting
    Input: $\gamma_R, \gamma_J > 0$ such that $\gamma_R\gamma_J\|L\|^2 < 1$ and $\tau \in [0,1]$. Let $s \geq 1, q \geq 1$ be integers.
    Initial: $\bar{x}_0 = x_0 \in \mathbb{R}^n$, $\bar{w}_0 = w_0 \in \mathbb{R}^m$. Let $z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix}$ and $V_0 \in \mathbb{R}^{(m+n)\times q}$.
    Repeat:
        • If mod$(k, q+2) = 0$: Compute $C_k$ via (5.1), if $\rho(C_k) < 1$: $e_k = \mathscr{E}_{s,q}(z_k, \cdots, z_{k-q-1})$.
              $\bar{x}_k = x_k + a_k e_{k,(1:n)}$,
              $\bar{w}_k = w_k + a_k e_{k,(n+1:m+n)}$.
        • If mod$(k, q+2) \neq 0$: $\bar{x}_k = x_k$ and $\bar{w}_k = w_k$.
        • For $k \geq 1$:
              $x_{k+1} = \mathrm{prox}_{\gamma_R R}(\bar{x}_k - \gamma_R L^T \bar{w}_k)$,
              $\tilde{x}_{k+1} = x_{k+1} + \tau(x_{k+1} - \bar{x}_k)$,
              $w_{k+1} = \mathrm{prox}_{\gamma_J J^*}(\bar{w}_k + \gamma_J L\tilde{x}_{k+1})$,
              $z_{k+1} = \begin{pmatrix} x_{k+1} \\ w_{k+1} \end{pmatrix}$,   $v_{k+1} = z_{k+1} - z_k$   and   $V_{k+1} = [v_{k+1}|v_k|\cdots|v_{k-q+2}]$.
    Until: $\|v_k\| \leq \mathrm{tol}$.
```

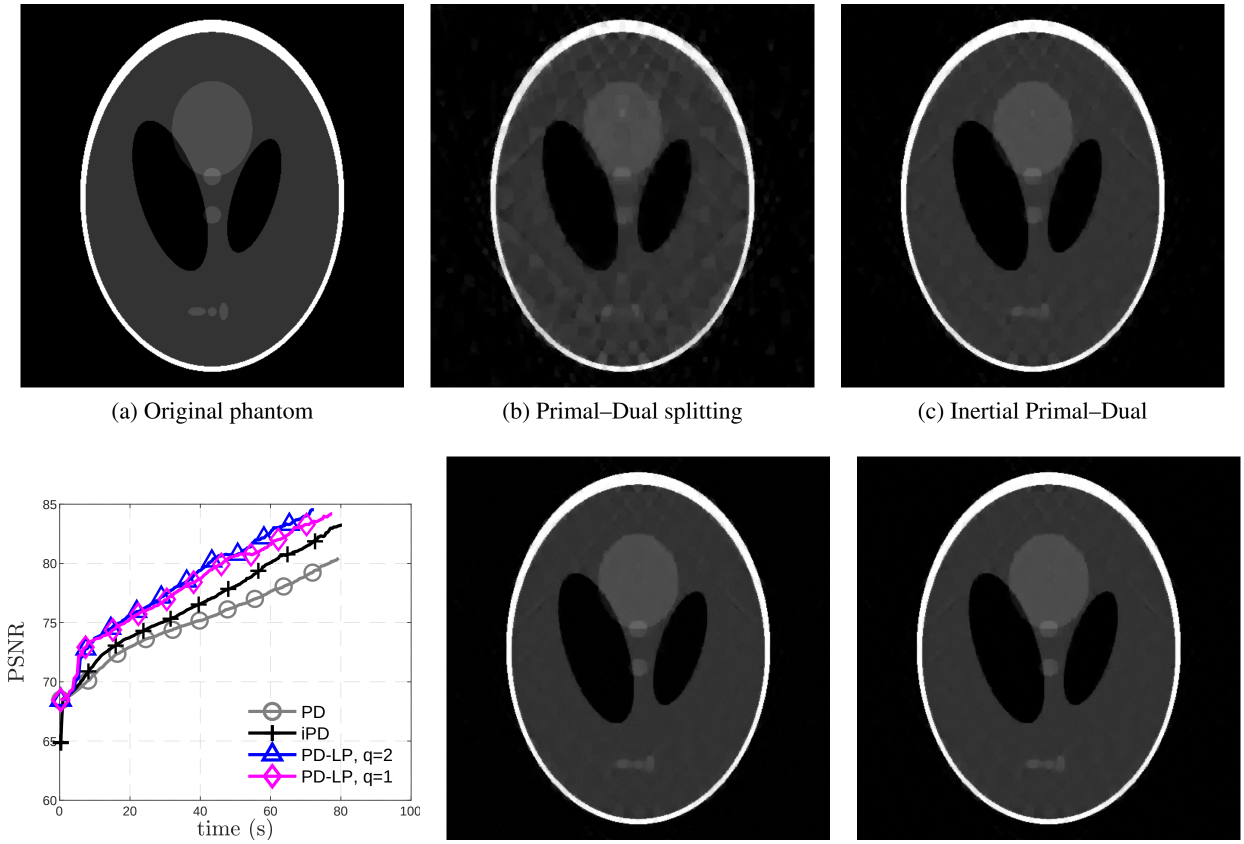
To demonstrate the performance of the above acceleration scheme, a medical imaging problem of the following form is considered

$$\min_{x\in\mathbb{R}^n} \|\mathscr{W}x\|_1 + \frac{\lambda}{2}\|\mathscr{K}x - f\|^2,$$

where $\mathscr{K}$ is a subsampled Fourier transform operator, f is the measurement and $\mathscr{W}$ a redundant wavelet frame. We compare the standard Primal–Dual splitting, inertial Primal–Dual splitting and our proposed accelerated one with $(q,s) = (1,+\infty), (2,+\infty)$, The numerical result is shown in Figure 13.

- Image quality wise, LP provides much better reconstruction than the plain Primal–Dual splitting and its inertial version, especially for $q = 2$.

- In terms of PSNR in Figure 13 (d), LP also yields better PSNR value than the (inertial) Primal–Dual splitting methods.



(a) Original phantom (b) Primal–Dual splitting (c) Inertial Primal–Dual

(d) PSNR (e) LP, $q = 1$ (f) LP, $q = 2$

Figure 13: Comparison of Primal–Dual schemes for MRI reconstruction. (d) Original Shepp–Logan phantom; (b) Output of Primal–Dual splitting; (c) Output of inertial Primal–Dual splitting; (d) PSNR comparison; (e) Output of Algorithm 11 with $(q,s) = (1,+\infty)$; (c) Output of Algorithm 11 with $(q,s) = (2,+\infty)$.

7.4 Generalized Forward–Backward splitting

For the problem ($\mathscr{P}_{\mathrm{FB}}$), suppose now there are more than 1 non-smooth functionals: let r be a positive integer and consider

$$\min_{x\in\mathbb{R}^n} \big\{ \Phi_r(x) \stackrel{\mathrm{def}}{=} F(x) + \textstyle\sum_{i=1}^{r} R_i(x) \big\}, \tag{$\mathscr{P}_{\mathrm{GFB}}$}$$

where F is continuous differentiable with ∇F being L-Lipschitz and $R_i \in \Gamma_0(\mathbb{R}^n)$ for each $i = 1, ..., r$.

Forward–Backward splitting is no longer feasible for this problem, as in general there is no close form solution for the proximity mapping of $\sum_{i=1}^{r} R_i(x)$ even if each R_i is simple. In [67], the authors proposed a generalized Forward–Backward splitting algorithm (GFB) to overcome the challenge. GFB achieves the full splitting of the evaluation of the proximity operator of each R_i. Let $(\omega_i)_i \in]0,1[^r$ such that $\sum_{i=1}^{r} \omega_i = 1$, choose $\gamma \in]0, 2\beta[$:

$$\begin{aligned}
&\text{from } i = 1 \text{ to } r: \\
&\left\lfloor \begin{aligned} u_{i,k+1} &= \mathrm{prox}_{\frac{\gamma}{\omega_i} R_i}\big(2x_k - z_{i,k} - \gamma\nabla F(x_k)\big) \\ z_{i,k+1} &= z_{i,k} + (u_{i,k+1} - x_k) \end{aligned} \right. \\
&x_{k+1} = \textstyle\sum_{i=1}^{r} \omega_i z_{i,k+1}.
\end{aligned} \tag{7.4}$$

Under a properly defined product space $\mathscr{H}$, there exists a non-expansive operator $\mathscr{F}_{\mathrm{GFB}} : \mathscr{H} \to \mathscr{H}$ such that

$$\boldsymbol{z}_{k+1} = \mathscr{F}_{\mathrm{GFB}}(\boldsymbol{z}_k)$$

with $\boldsymbol{z}_k = \begin{pmatrix} z_{1,k} \\ \vdots \\ z_{r,k} \end{pmatrix}$. We refer to [67] for more details of the GFB algorithm.

Specializing A^{2}FoM to the case of GFB, we obtain the following accelerated GFB scheme.

```
**Algorithm 12:** A$^2$GFB : Adaptive Acceleration for generalized Forward–Backward splitting
    **Input:** $(\omega_j)_j \in ]0,1[^r$ such that $\sum_{j=1}^r \omega_j = 1$, $\gamma \in ]0, 2/L[$. Let $s \geq 1, q \geq 1$ be integers.
    **Initial**: for $i = 1, ..., r$, $\bar{z}_{i,0} = z_0 \in \mathbb{R}^n$ and $x_0 = \sum_{i=1}^r \omega_i \bar{z}_{i,0}$, $V_{i,0} = 0 \in \mathbb{R}^{n \times q}$;
    **Repeat**:
      • If $\mathrm{mod}(k, q+2) = 0$: for $i = 1, ..., r$, compute $C_k^i$ via (5.1), if $\rho(C_k^i) < 1$:
            $\bar{z}_{i,k} = z_{i,k} + a_k^i \mathscr{E}_{s,q}(z_{i,k}, \cdots, z_{i,k-q-1})$.
      • If $\mathrm{mod}(k, q+2) \neq 0$:  $\bar{z}_{i,k} = z_{i,k}$.
      • For $k \geq 1$:
            $x_k = \sum_{i=1}^r \omega_i \bar{z}_{i,k}$,
            from $i = 1$ to $r$:
              $u_{i,k+1} = \mathrm{prox}_{\frac{\gamma}{\omega_i} R_i}\big(2x_k - \bar{z}_{i,k} - \gamma \nabla F(x_k)\big)$,
              $z_{i,k+1} = \bar{z}_{i,k} + (u_{i,k+1} - x_k)$,
              $v_{i,k+1} = z_{i,k+1} - z_{i,k}$  and  $V_{i,k+1} = [v_{i,k+1} | v_{i,k} | \cdots | v_{i,k-q+2}]$.
    **Until**: $\sum_i \|v_k^i\| \leq \mathrm{tol}$.
```

We consider the Principal Component Pursuit (PCP) problem [22] to demonstrate the performance comparison. Different from (7.2), the forward observation model of PCP problem reads,

$$b = \mathring{x}_{L} + \mathring{x}_{S} + \omega,$$

where $\mathring{x}_{L}$ is low-rank, $\mathring{x}_{S}$ is sparse, and b, ω are the observation and noise respectively. The PCP proposed in [22] attempts to provably recover $(\mathring{x}_{L}, \mathring{x}_{S})$ up to a good approximation, by solving a convex optimization. Here, we also add a non-negativity constraint to the low-rank component, which leads to the following convex problem

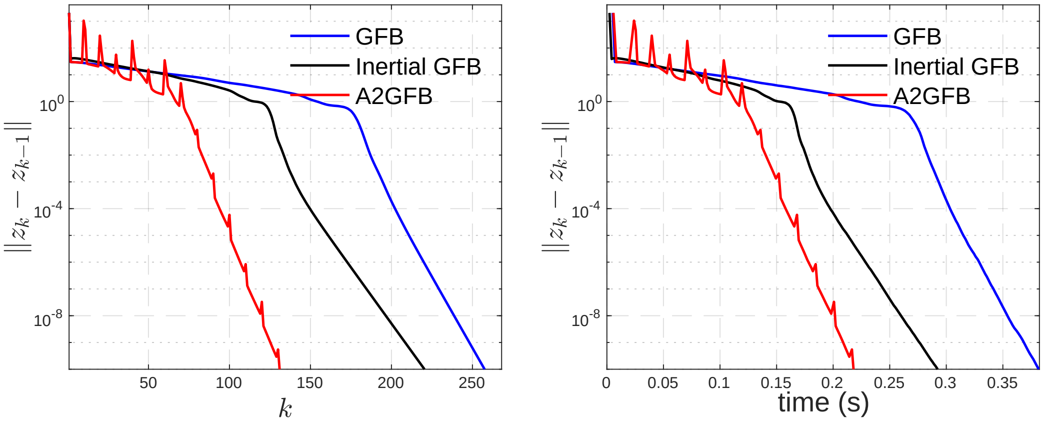
$$\min_{x_{L}, x_{S} \in \mathbb{R}^{n \times n}} \frac{1}{2}\|b - x_{L} - x_{S}\|^2 + \mu_1 \|x_{S}\|_1 + \mu_2 \|x_{L}\|_* + \iota_{P_+}(x_{L}). \tag{7.5}$$

Observe that for given an x_{L}, the minimizer of (7.5) is $x_{S}^\star = \mathrm{prox}_{\mu_1 \|\cdot\|_1}(b - x_{L})$. Thus, (7.5) is equivalent to

$$\min_{x_{L} \in \mathbb{R}^{n \times n}} {}^1\big(\mu_1 \|\cdot\|_1\big)(b - x_{L}) + \mu_2 \|x_{L}\|_* + \iota_{P_+}(x_{L}), \tag{7.6}$$

where ${}^1\big(\mu_1 \|\cdot\|_1\big)(b - x_{L})$ is the Moreau Envelope of $\mu_1 \|\cdot\|_1$.

The numerical comparison on a synthetic example is shown below in Figure 14, and the observations are very similar to those of the previous examples, that linear prediction shows clear advantages over the standard method and its inertial version.



(a) Iteration comparison

(b) CPU time

Figure 14: Comparison of GFB and adaptive accelerated GFB on synthetic data.

8 Conclusions

In this article, by analyzing the trajectory of the fixed point sequences associated to first-order methods and extrapolating along the trajectory, we provide an alternative derivation of these methods. Furthermore, our local linear analysis allows for the application of previous results on extrapolation methods, and hence provides guaranteed (local) acceleration.

Acknowledgement

We would like to thank Arieh Iserles for pointing out the connection between linear prediction and vector extrapolation. Jingwei Liang was partly supported by Leverhulme trust, Newton trust, the EPSRC centre “EP/N014588/1” and the Cantab Capital Institute for the Mathematics of Information (CCIMI).

References

[1] P-A. Absil, R. Mahony, and J. Trumpf. An extrinsic look at the Riemannian Hessian. In *Geometric Science of Information*, pages 361–368. Springer, 2013.

[2] A. C. Aitken. Xxv.–on Bernoulli’s numerical solution of algebraic equations. *Proceedings of the Royal Society of Edinburgh*, 46:289–305, 1927.

[3] F. Alvarez. On the minimizing property of a second order dissipative system in Hilbert spaces. *SIAM Journal on Control and Optimization*, 38(4):1102–1119, 2000.

[4] F. Alvarez and H. Attouch. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Analysis*, 9(1-2):3–11, 2001.

[5] D. G. Anderson. Iterative procedures for nonlinear integral equations. *J. ACM*, 12(4):547–560, October 1965.

[6] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized gauss–seidel methods. *Mathematical Programming*, 137(1-2):91–129, 2013.

[7] H. Bauschke and P. L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2011.

[8] H. H. Bauschke, J. Y. Bello Cruz, T. T. A. Nghia, H. M. Pha, and X. Wang. Optimal rates of linear convergence of relaxed alternating projections and generalized Douglas–Rachford methods for two subspaces. *Numerical Algorithms*, 73(1):33–76, 2016.

[9] H. H. Bauschke, J. Y. Bello Cruz, T. T. Nghia, H. M. Pha, and X. Wang. Optimal rates of linear convergence of relaxed alternating projections and generalized douglas-rachford methods for two subspaces. *Numerical Algorithms*, 73(1):33–76, 2016.

[10] H. H. Bauschke, JY B. Cruz, T. TA Nghia, H. M. Phan, and X. Wang. The rate of linear convergence of the douglas–rachford algorithm for subspaces is the cosine of the friedrichs angle. *Journal of Approximation Theory*, 185:63–79, 2014.

[11] 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.

[12] R. Bollapragada, D. Scieur, and A. d'Aspremont. Nonlinear acceleration of momentum and primal-dual algorithms. *arXiv preprint arXiv:1810.04539*, 2018.

[13] P. Borwein, C. Pinner, and I. Pritsker. Monic integer chebyshev problem. *Mathematics of computation*, 72(244):1901–1916, 2003.

[14] R. I. Boţ and E. Csetnek. An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems. *Numerical Algorithms*, 71(3):519–540, 2016.

[15] R. I. Bot and E. R. Csetnek. An inertial alternating direction method of multipliers. *arXiv preprint arXiv:1404.4582*, 2014.

[16] R. I. Boţ, E. R. Csetnek, and C. Hendrich. Inertial Douglas–Rachford splitting for monotone inclusion problems. *Applied Mathematics and Computation*, 256:472–487, 2015.

[17] C. Brezinski. Convergence acceleration during the 20th century. *Numerical Analysis: Historical Developments in the 20th Century*, page 113, 2001.

[18] C. Brezinski, M. Redivo-Zaglia, and Y. Saad. Shanks sequence transformations and anderson acceleration. *SIAM Review*, 60(3):646–669, 2018.

[19] C. Brezinski and M. R. Zaglia. *Extrapolation methods: theory and practice*, volume 2. Elsevier, 2013.

[20] L. M. Briceno-Arias and P. L. Combettes. A monotone+ skew splitting model for composite monotone inclusions in duality. *SIAM Journal on Optimization*, 21(4):1230–1250, 2011.

[21] S. Cabay and L. W. Jackson. A polynomial extrapolation method for finding limits and antilimits of vector sequences. *SIAM Journal on Numerical Analysis*, 13(5):734–752, 1976.

[22] E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011.

[23] A. Chambolle and C. Dossal. On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm". *Journal of Optimization Theory and Applications*, 166(3):968–982, 2015.

[24] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[25] R. H. Chan, S. Ma, and J. Yang. Inertial primal-dual algorithms for structured convex optimization. *arXiv preprint arXiv:1409.2992*, 2014.

[26] I. Chavel. *Riemannian geometry: a modern introduction*, volume 98. Cambridge University Press, 2006.

[27] P. L. Combettes, L. Condat, J.-C. Pesquet, and B. C. Vũ. A Forward–Backward view of some Primal–Dual optimization methods in image recovery. In *Image Processing (ICIP), 2014 IEEE International Conference on*, pages 4141–4145. IEEE, 2014.

[28] P. L. Combettes and B. C. Vũ. Variable metric Forward–Backward splitting with applications to monotone inclusions in duality. *Optimization*, 63(9):1289–1318, 2014.

[29] L. Demanet and X. Zhang. Eventual linear convergence of the douglas-rachford iteration for basis pursuit. *Mathematics of Computation*, 85(297):209–238, 2016.

[30] Q. Dong, Yeol J. Cho, and T. M. Rassias. General inertial mann algorithms and their convergence analysis for nonexpansive mappings. In *Applications of Nonlinear Analysis*, pages 175–191. Springer, 2018.

[31] Q. Dong, J. Huang, X. Li, Y. Cho, and T. M. Rassias. Mikm: multi-step inertial krasnosel'skiĭ–mann algorithm and its applications. *Journal of Global Optimization*, 73(4):801–824, 2019.

[32] Q. Dong, H. Yuan, Y. Cho, and T. M. Rassias. Modified inertial mann algorithm and inertial cq-algorithm for nonexpansive mappings. *Optimization Letters*, 12(1):87–102, 2018.

[33] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American mathematical Society*, 82(2):421–439, 1956.

[34] R. P. Eddy. Extrapolating to the limit of a vector sequence. In *Information linkage between applied mathematics and industry*, pages 387–396. Elsevier, 1979.

[35] J. Fadili, J. Malick, and G. Peyré. Sensitivity analysis for mirror-stratifiable convex functions. *SIAM Journal on Optimization*, 28(4):2975–3000, 2018.

[36] G. França, D. P. Robinson, and R. Vidal. Admm and accelerated admm as continuous dynamical systems. *arXiv preprint arXiv:1805.06579*, 2018.

[37] D. Gabay and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.

[38] L. G. Gubin, B. T. Polyak, and E. V. Raik. The method of projections for finding the common point of convex sets. *USSR Computational Mathematics and Mathematical Physics*, 7(6):1–24, 1967.

[39] W. L. Hare and A. S. Lewis. Identifying active constraints via partial smoothness and prox-regularity. *Journal of Convex Analysis*, 11(2):251–266, 2004.

[40] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 1990.

[41] M. Kadkhodaie, K. Christakopoulou, M. Sanjabi, and A. Banerjee. Accelerated alternating direction method of multipliers. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pages 497–506. ACM, 2015.

[42] M. A. Krasnosel'skii. Two remarks on the method of successive approximations. *Uspekhi Matematicheskikh Nauk*, 10(1):123–127, 1955.

[43] J. M. Lee. *Smooth manifolds*. Springer, 2003.

[44] A. S. Lewis. Active sets, nonsmoothness, and sensitivity. *SIAM Journal on Optimization*, 13(3):702–725, 2003.

[45] J. Liang. *Convergence rates of first-order operator splitting methods*. PhD thesis, Normandie Université; GREYC CNRS UMR 6072, 2016.

[46] J. Liang, J. Fadili, and G. Peyré. Local linear convergence of Forward–Backward under partial smoothness. In *Advances in Neural Information Processing Systems*, pages 1970–1978, 2014.

[47] J. Liang, J. Fadili, and G. Peyré. Activity identification and local linear convergence of Forward–Backward-type methods. *SIAM Journal on Optimization*, 27(1):408–437, 2017.

[48] J. Liang, J. Fadili, and G. Peyré. Local convergence properties of Douglas–Rachford and alternating direction method of multipliers. *Journal of Optimization Theory and Applications*, 172(3):874–913, 2017.

[49] J. Liang, J. Fadili, and G. Peyré. Local linear convergence analysis of primal–dual splitting methods. *Optimization*, 67(6):821–853, 2018.

[50] J. Liang, T. Luo, and C. Schönlieb. Improving "fast iterative shrinkage-thresholding algorithm": Faster, smarter and greedier. *arXiv preprint arXiv:1811.01430*, 2018.

[51] P. L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.

[52] D. A. Lorenz and T. Pock. An inertial forward-backward algorithm for monotone inclusions. *Journal of Mathematical Imaging and Vision*, 51(2):311–325, 2015.

[53] P. Maingé. Convergence theorems for inertial km-type algorithms. *Journal of Computational and Applied Mathematics*, 219(1):223–236, 2008.

[54] W. R. Mann. Mean value methods in iteration. *Proceedings of the American Mathematical Society*, 4(3):506–510, 1953.

[55] M. Mešina. Convergence acceleration for the iterative solution of the equations x= ax+ f. *Computer Methods in Applied Mechanics and Engineering*, 10(2):165–173, 1977.

[56] S. A. Miller and J. Malick. Newton methods for nonsmooth convex minimization: connections among-Lagrangian, Riemannian Newton and SQP methods. *Mathematical programming*, 104(2-3):609–633, 2005.

[57] C. Molinari, J. Liang, and J. Fadili. Convergence rates of forward–douglas–rachford splitting method. *arXiv preprint arXiv:1801.01088*, 2018.

[58] A. Moudafi and M. Oliny. Convergence of a splitting inertial proximal method for monotone operators. *Journal of Computational and Applied Mathematics*, 155(2):447–454, 2003.

[59] Y. Nesterov. A method for solving the convex programming problem with convergence rate $O(1/k^2)$. *Dokl. Akad. Nauk SSSR*, 269(3):543–547, 1983.

[60] B. O'Donoghue and E. Candes. Adaptive restart for accelerated gradient schemes. *Foundations of computational mathematics*, 15(3):715–732, 2015.

[61] M. Özdemir. An alternative approach to elliptical motion. *Advances in Applied Clifford Algebras*, 26(1):279–304, 2016.

[62] I. Pejcic and C. N. Jones. Accelerated admm based on accelerated douglas-rachford splitting. In *2016 European Control Conference (ECC)*, pages 1952–1957. Ieee, 2016.

[63] Y. Peng, B. Deng, J. Zhang, F. Geng, W. Qin, and L. Liu. Anderson acceleration for geometry optimization and physics simulation. *ACM Transactions on Graphics (TOG)*, 37(4):1–14, 2018.

[64] B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964.

[65] B. T. Polyak. *Introduction to optimization*. Optimization Software, 1987.

[66] C. Poon and J. Liang. Trajectory of alternating direction method of multipliers and adaptive acceleration. In *Advances In Neural Information Processing Systems*, 2019.

[67] H. Raguet, M. J. Fadili, and G. Peyré. Generalized forward-backward splitting. *SIAM Journal on Imaging Sciences*, 6(3):1199–1226, 2013.

[68] L. F. Richardson and J. A. Gaunt. Viii. the deferred approach to the limit. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, 226(636-646):299–361, 1927.

[69] R. T. Rockafellar. *Convex analysis*, volume 28. Princeton university press, 1997.

[70] D. Scieur, F. Bach, and A. d'Aspremont. Nonlinear acceleration of stochastic algorithms. In *Advances in Neural Information Processing Systems*, pages 3982–3991, 2017.

[71] D. Scieur, A. d'Aspremont, and F. Bach. Regularized nonlinear acceleration. In *Advances In Neural Information Processing Systems*, pages 712–720, 2016.

[72] D. Shanks. Non-linear transformations of divergent and slowly convergent sequences. *Journal of Mathematics and Physics*, 34(1-4):1–42, 1955.

[73] A. Sidi. *Practical extrapolation methods: Theory and applications*, volume 10. Cambridge University Press, 2003.

[74] A. Sidi. Vector extrapolation methods with applications to solution of large systems of equations and to pagerank computations. *Computers & Mathematics with Applications*, 56(1):1–24, 2008.

[75] A. Sidi. *Vector extrapolation methods with applications*, volume 17. SIAM, 2017.

[76] W. Su, S. Boyd, and E. Candes. A differential equation for modeling nesterov's accelerated gradient method: Theory and insights. In *Advances in Neural Information Processing Systems*, pages 2510–2518, 2014.

[77] S. Vaiter, G. Peyré, and J. Fadili. Model consistency of partly smooth regularizers. *IEEE Transactions on Information Theory*, 64(3):1725–1737, 2018.

[78] B. C. Vũ. A splitting algorithm for dual monotone inclusions involving cocoercive operators. *Advances in Computational Mathematics*, 38(3):667–681, 2013.

[79] P. Wynn. Acceleration techniques for iterated vector and matrix problems. *Mathematics of Computation*, 16(79):301–322, 1962.

[80] J. Zhang, B. O'Donoghue, and S. Boyd. Globally convergent type-i anderson acceleration for non-smooth fixed-point iterations. *arXiv preprint arXiv:1808.03971*, 2018.

A Trajectory of linear systems

Let $M \in \mathbb{R}^{n\times n}$ be a bounded real matrix and consider the following linear system

$$x_{k+1} = Mx_k, \tag{A.1}$$

which generates a train of sequence $\{x_k\}_{k\in\mathbb{N}}$. Assumed $\{x_k\}_{k\in\mathbb{N}}$ is convergent, *i.e.* there exists an $x^\star \in \mathbb{R}^n$ such that $x_k \to x^\star$. The goal of this section is to investigate the properties of the trajectory formed by $\{x_k\}_{k\in\mathbb{N}}$. To this end, define $v_k = x_k - x_{k-1}$, it is immediate that (A.1) leads to the following iteration in terms of v_k,

$$v_{k+1} = Mv_k, \tag{A.2}$$

and $\lim_{k\to+\infty} v_k = 0$ since $x_k \to x^\star$. To characterize the trajectory, we choose to use the angle between each two adjacent vectors v_k and v_{k-1}, which is denoted by θ_k and defined by

$$\theta_k \overset{\text{def}}{=} \angle(v_k, v_{k-1}) = \arccos\Big(\frac{\langle v_k,\, v_{k-1}\rangle}{\|v_k\|\|v_{k-1}\|}\Big).$$

For the rest of this section, we discuss the property of $\{\theta_k\}_{k\in\mathbb{N}}$ under four different choices of matrix M.

A.1 Type I linear system

We start with the simplest case, that $M \in \mathbb{R}^{n\times n}$ is symmetric. Let $(\sigma_i)_{i=1,\ldots,n} \in \mathbb{R}^n$ be the eigenvalues of M, which are all real owing to the symmetry of M.

Definition A.1 (Type I matrix). $M \in \mathbb{R}^{n\times n}$ is symmetric with all its eigenvalues in $]-1,1]$, moreover $1 \geq \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$ and $\sigma_1 > |\sigma_n| > 0$.

Denote η the ratio between the second largest eigenvalue in magnitude and σ_1, *i.e.* $\eta = \frac{\max\{\sigma_2, |\sigma_n|\}}{\sigma_1}$.

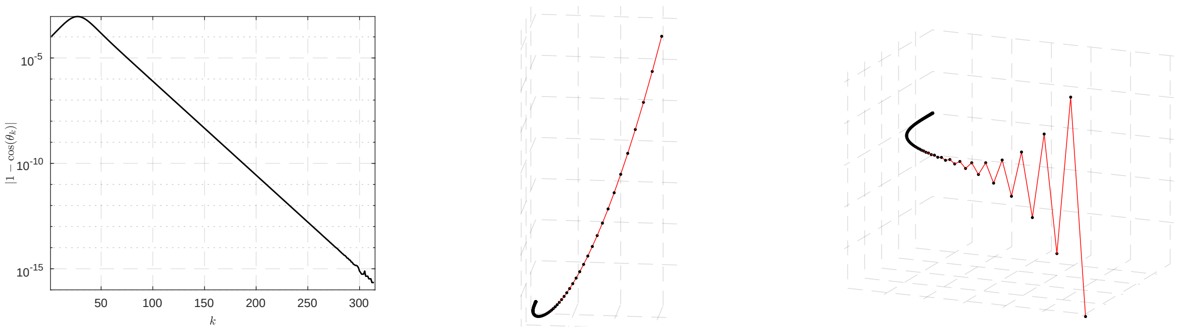
Proposition A.1. *Consider the linear system* (A.2) *where M is a Type I matrix defined in Definition A.1, then there holds* $1 - \cos(\theta_k) = O(\eta^{2k})$.

Remark A.2. Proposition A.1 implies eventually the trajectory of $\{x_k\}_{k\in\mathbb{N}}$ is a straight line. If $\sigma_1 < |\sigma_n| < 1$, then it can be shown that $\lim_{k\to+\infty}\theta_k = \pi$ and $\vartheta_k \to 0$.

Example A.2. Let U be an orthogonal matrix in $\mathbb{R}^{3\times 3}$, and consider $M = U\begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix}U^T$ where $-1 < c \leq b \leq a \leq 1$ are the eigenvalues. Two different choices of (a,b,c) are considered

$$(a,b,c) \in \big\{0.99\times(1, 0.98, 0.9), 0.99\times(1, 0.98, -0.75)\big\}.$$

For both cases, we have $\eta = 0.98$, hence same convergence rates of $\cos(\theta_k)$ to 1, see Figure 15 (a). The trajectories of $\{x_k\}_{k\in\mathbb{N}}$ are shown in the other two figures of Figure 15: for figure (b) all the three eigenvalues of M are in $[0,1]$, for figure (c) the smallest eigenvalue of M is negative.



(a) $1-\cos(\theta_k)$ (b) Eigenvalues in $[0,1]$ (c) Eigenvalues in $]-1,1]$

Figure 15: Convergence of $\cos(\theta_k)$ and trajectories of $\{x_k\}_{k\in\mathbb{N}}$. (a): Convergence of $\cos(\theta_k)$ to 1; (b) Trajectory of $\{x_k\}_{k\in\mathbb{N}}$ when all the eigenvalue of M are real; (c) Trajectory of $\{x_k\}_{k\in\mathbb{N}}$ when M has negative eigenvalues.

Proof of Proposition A.1. Since M is symmetric, there exists a real orthogonal matrix U such that $M = U\Sigma U^T$ where $\Sigma = \mathrm{diag}((\sigma_i)_{i=1,\dots,n})$ is a diagonal matrix, and $v_k = Mv_{k-1} = M^k v_0 = U\Sigma^k U^T v_0$. Let $u_k = U^T v_k$, then $u_k = \Sigma^k u_0$. Suppose there exists $d \in [2,n[$ such that $\sigma = \sigma_1 = \sigma_2 = \cdots = \sigma_d > \sigma_{d+1}$, we can consider the following decomposition of Σ

$$\Sigma_1 \overset{\text{def}}{=} \begin{bmatrix} \mathrm{diag}((\sigma_i)_{i=1,\dots,d}) & 0_{d\times(n-d)} \\ 0_{(n-d)\times d} & 0_{n-d} \end{bmatrix} \quad \text{and} \quad \Sigma_2 \overset{\text{def}}{=} \begin{bmatrix} 0_d & 0_{d\times(n-d)} \\ 0_{(n-d)\times d} & \mathrm{diag}((\sigma_i)_{i=d+1,\dots,n}) \end{bmatrix}. \tag{A.3}$$

It is immediate that $u_k = \Sigma_1^k u_0 + \Sigma_2^k u_0$, and that

$$\frac{1}{\sigma}\Sigma_1 = \begin{bmatrix} \mathrm{Id}_d & 0_{d\times(n-d)} \\ 0_{(n-d)\times d} & 0_{n-d} \end{bmatrix} \quad \text{and} \quad \frac{1}{\sigma}\Sigma_2 = \eta \begin{bmatrix} 0_d & 0_{d\times(n-d)} \\ 0_{(n-d)\times d} & \mathrm{diag}((\frac{\sigma_i}{\sigma_{d+1}})_{i=d+1,\dots,n}) \end{bmatrix}.$$

Moreover, there holds $\frac{1}{\sigma^k}\Sigma_1^k = \frac{1}{\sigma}\Sigma_1$ and $\frac{1}{\sigma^k}\Sigma_2^k = O(\eta^k)$. Consider the following orthogonal decomposition of $\frac{u_k}{\sigma^k}$,

$$s_k = \frac{1}{\sigma^k}\Sigma_1^k u_0 = \frac{1}{\sigma}\Sigma_1 u_0 \quad \text{and} \quad t_k = \frac{u_k}{\sigma^k} - s_k = O(\eta^k).$$

We get

$$\langle v_k,\, v_{k-1}\rangle = \langle U^T v_k,\, U^T v_{k-1}\rangle = \sigma^{2k-1}\langle \tfrac{u_k}{\sigma^k},\, \tfrac{u_{k-1}}{\sigma^{k-1}}\rangle = \sigma^{2k-1}\langle s_k + t_k,\, s_{k-1} + t_{k-1}\rangle,$$

and $\|v_k\| = \|u_k\| = \sigma^k(\|s_k + t_k\|) = \sigma^k(\|s_k\| + \|t_k\|)$. Consequently the value of $\cos(\theta_k)$ is, note that $s_k = s_{k-1}$

$$\begin{aligned} \cos(\theta_k) = \frac{\langle v_k,\, v_{k-1}\rangle}{\|v_k\|\|v_{k-1}\|} = \frac{\langle s_k + t_k,\, s_{k-1} + t_{k-1}\rangle}{\|s_k + t_k\|\|s_{k-1} + t_{k-1}\|} &= \frac{\langle s_k,\, s_{k-1}\rangle}{\|s_k + t_k\|\|s_{k-1} + t_{k-1}\|} + \frac{\langle t_k,\, t_{k-1}\rangle}{\|s_k + t_k\|\|s_{k-1} + t_{k-1}\|} \\ &= \frac{\|s_k\|^2}{\|s_k + t_k\|\|s_{k-1} + t_{k-1}\|} + O(\eta^{2k-1}) \\ &= \frac{\|s_k\|^2}{\|s_k\|^2 + \|t_k\|^2} \times \frac{\|s_k + t_k\|}{\|s_k + t_{k-1}\|} + O(\eta^{2k-1}). \end{aligned} \tag{A.4}$$

Since we have

$$\frac{\|s_k\|^2}{\|s_k\|^2 + \|t_k\|^2} = 1 - \|t_k\|^2 + O(\|t_k\|^4) = 1 + O(\eta^{2k}) \quad \text{and} \quad \frac{\|s_k + t_k\|}{\|s_k + t_{k-1}\|} \to 1.$$

Combining with (A.4) leads to the claimed result. □

A.2 Type II linear system

From this part, we turn to linear systems which result in spiral trajectories. The first of this kind is the normal matrix.

Definition A.3 (Type II matrix). $M \in \mathbb{R}^{n\times n}$ is normal matrix with all its eigenvalues lying in the complex unit disc.

According to [40, Theorem 2.5.8], a normal matrix $M \in \mathbb{R}^{n\times n}$ is quasi-diagonalizable, that is there exists a real orthogonal matrix $U \in \mathbb{R}^{n\times n}$ such that

$$M = U \begin{bmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{bmatrix} U^T.$$

For each $i = 1,\dots,m$, B_i is either real valued scalar or 2×2 matrix of the form $\begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}$ in which $b_i > 0$ and has eigenvalues $a_i \pm \mathrm{i}b_i$. We impose the following assumptions on M.

Assumption A.3. *For each $i = 1,\dots,m$,*

(i) *if B_i is scalar, then $B_i \in \{0,1\}$;*

(ii) *if $B_i = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}$ with $b_i > 0$, then $a_i^2 + b_i^2 < 1$. Moreover, there exists $1 \le q \le d \le m$ such that $B_1 = \cdots = B_q$ and $1 > a_1^2 + b_1^2 = a_2^2 + b_2^2 = \cdots = a_q^2 + b_q^2 > a_{q+1}^2 + b_{q+1}^2 \ge \cdots \ge a_d^2 + b_d^2 > 0$.*

Let ψ be the argument of $a_1 + \mathrm{i}b_1$ and $\eta = \frac{\sqrt{a_{q+1}^2+b_{q+1}^2}}{\sqrt{a_q^2+b_q^2}}$. Under Assumption A.3, the power of M, *i.e.* M^k, is convergent when k goes to $+\infty$. Denote $M^\infty \overset{\text{def}}{=} \lim_{k\to+\infty} M^k$.

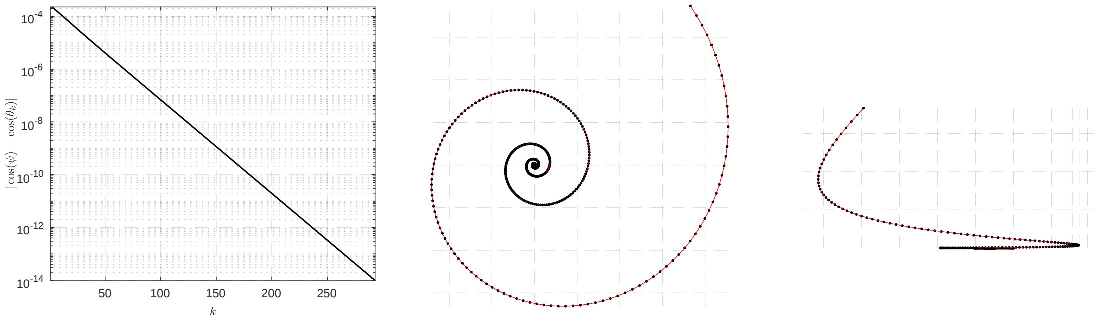
Proposition A.4. *Consider the linear system (A.2) whose M is a Type II matrix define in Definition A.3, suppose that Assumption A.3 holds. Then*

(i) *M^∞ is a symmetric matrix with eigenvalues being either 0 or 1, and $v_0 \in \ker(M^\infty)$.*

(ii) *$\cos(\theta_k) - \cos(\psi) = O(\eta^{2k})$ for ψ and η defined above.*

Remark A.5.

- Proposition A.4 indicates that eventually M performs circular rotation.
- If we have only $a_1^2 + b_1^2 = a_2^2 + b_2^2 = \cdots = a_q^2 + b_q^2$, and $B_i \ne B_j, 1 \le i,j \le q, i \ne j$, then $\cos(\theta_k)$ will converge to some ψ which depends on $(\psi_i)_{i=1,\dots,q}$ where ψ_i is the argument of $a_i + \mathrm{i}b_i$.



(a) $|\cos(\theta_k) - \cos(\psi)|$ (b) Top view (c) Side view

Figure 16: Convergence of $\cos(\theta_k)$ and trajectory of $\{x_k\}_{k\in\mathbb{N}}$. (a): Convergence of $\cos(\theta_k)$ to $\cos(\psi)$; (b) Top view of trajectory of $\{x_k\}_{k\in\mathbb{N}}$; (c) Side view of trajectory of $\{x_k\}_{k\in\mathbb{N}}$.

Example A.4. Let $\alpha \in]0,\pi/2]$ and define a,b,c by

$$a = 0.99\cos(\alpha),\;\; b = 0.99\sin(\alpha) \quad\text{and}\quad c = \eta\sqrt{a^2+b^2}$$

for some $\eta \in]0,1[$. Let U be an orthogonal matrix in $\mathbb{R}^{3\times 3}$, and let $M = U \begin{bmatrix} a & b & \\ -b & a & \\ & & c \end{bmatrix} U^T$. M has three eigenvalues: $a + \mathrm{i}b, a - \mathrm{i}b$ and c. Let ψ be the argument of $a + \mathrm{i}b$ and $(\alpha,\eta) = (0.05, 0.96)$, The convergence of $\cos(\theta_k)$ and trajectories of $\{x_k\}_{k\in\mathbb{N}}$ are provided in Figure 16. The first plot shows the convergence of $\cos(\theta_k)$ to $\cos(\phi)$, and the other two are different views of the trajectory of $\{x_k\}_{k\in\mathbb{N}}$.

Proof of Proposition A.4. Owing to [40, Theorem 2.5.8] and Assumption A.3, we have the decomposition of M

$$M = U\Sigma U^T \quad \text{with} \quad \Sigma \overset{\text{def}}{=} \begin{bmatrix} \mathrm{Id}_r & & & & \\ & B_1 & & & \\ & & \ddots & & \\ & & & B_d & \\ & & & & 0_r \end{bmatrix}$$

where r denotes the multiplicities of eigenvalue 1 in A.3 (i), and d denotes the number of 2×2 blocks. For each $i = 1,...,d$, we have $B_i = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}$ with $1 > a_1^2+b_1^2 \geq a_2^2+b_2^2 \geq \cdots \geq a_d^2+b_d^2 > 0$. It is easy to show that $\lim_{k\to+\infty} B_i^k = 0, i = 1,...,d$ since the spectral radius of each B_i, $\rho(B_i) = \sqrt{a_i^2+b_i^2} < 1$, is strictly smaller than 1. This further implies that

$$M^\infty \overset{\text{def}}{=} \lim_{k\to+\infty} M^k = U \begin{bmatrix} \mathrm{Id}_r & \\ & 0_{n-r} \end{bmatrix} U^T,$$

which verifies the first claim of the proposition.

Since $v_k \to 0$, we have from $v_k = Mv_{k-1} = M^k v_0$ that

$$0 = \lim_{k\to+\infty} M^k v_0 = M^\infty v_0,$$

which means $v_0 \in \ker(M^\infty)$ and moreover $v_k \in \ker(M^\infty), k \in \mathbb{N}$. Consequently, we have

$$v_k = Mv_{k-1} = (M - M^\infty)v_{k-1}$$

Define $\widetilde{M} \overset{\text{def}}{=} M - M^\infty$, then there exists a real orthogonal matrix V (actually a permutation of U) such that

$$\widetilde{M} = V\Gamma V^T \quad \text{with} \quad \Gamma \overset{\text{def}}{=} \begin{bmatrix} B_1 & & & \\ & \ddots & & \\ & & B_d & \\ & & & 0_{2r} \end{bmatrix}$$

and $B_i = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}, i = 1,...,d$. Suppose for some $1 \leq q < d$, there holds

$$a_1^2+b_1^2 = a_2^2+b_2^2 = \cdots = a_q^2+b_q^2 > a_{q+1}^2+b_{q+1}^2 \geq \cdots \geq a_d^2+b_d^2.$$

Consider the decomposition of Γ,

$$\Gamma_1 = \begin{bmatrix} B_1 & & & \\ & \ddots & & \\ & & B_q & \\ & & & 0_{n-2q} \end{bmatrix} \quad \text{and} \quad \Gamma_2 = \Gamma - \Gamma_1.$$

Let $\sigma = \sqrt{a_1^2+b_1^2}$ and $\eta = \frac{\sqrt{a_{q+1}^2+b_{q+1}^2}}{\sigma}$, then $\frac{1}{\sigma^k}\Gamma_2^k = O(\eta^k) \to 0$. Let $\psi = \arccos(\frac{a_1}{\sigma})$, then for each $i = 1,...,q$

$$\frac{1}{\sigma}B_i = \begin{bmatrix} \cos(\psi) & \sin(\psi) \\ -\sin(\psi) & \cos(\psi) \end{bmatrix}$$

which is a circular rotation. Therefore, $\frac{1}{\sigma}\Gamma_1$ is a rotation with respect to the first $2q$ elements. Denote $u_k = V^T v_k$, then from $v_k = \widetilde{M}v_{k-1}$, we get $u_k = \Gamma u_{k-1} = \Gamma^k u_0$. Consider the orthogonal decomposition of $\frac{u_k}{\sigma^k}$,

$$s_k = \frac{1}{\sigma^k}\Gamma_1^k u_0 \quad \text{and} \quad t_k = \frac{1}{\sigma^k}\Gamma_2^k u_0.$$

We have that $\|s_k\| = \|s_{k-1}\|$ and $\langle s_k, s_{k-1}\rangle = \|s_k\|^2 \cos(\psi)$. As a result, for $\cos(\theta_k)$ we have

$$\cos(\theta_k) = \frac{\langle s_k, s_{k-1}\rangle}{\|s_k+t_k\|\|s_{k-1}+t_{k-1}\|} + \frac{\langle t_k, t_{k-1}\rangle}{\|s_k+t_k\|\|s_{k-1}+t_{k-1}\|} = \frac{\|s_k\|^2\cos(\psi)}{\|s_k\|^2+\|t_k\|^2} \times \frac{\|s_k+t_k\|}{\|s_{k-1}+t_{k-1}\|} + O(\eta^{2k-1}). \tag{A.5}$$

Using the fact that $\frac{\|s_k\|^2\cos(\psi)}{\|s_k\|^2+\|t_k\|^2} = \cos(\psi)\big(1-\|t_k\|^2+O(\|t_k\|^4)\big) = \cos(\psi)+O(\eta^{2k})$ and $\frac{\|s_k+t_k\|}{\|s_{k-1}+t_{k-1}\|} \to 1$ we conclude the convergence of θ_k. □

A.3 Type III linear system

The last trajectory we discuss is elliptical spiral which is more complicated. We first discuss the definition and properties of elliptical rotation, then discuss two types of matrices that lead to elliptical rotation.

A.3.1 Elliptical rotation

Definition A.5 ([61, Theorem 1]). Let $\phi > 0$ and $l > s > 0$, then the following matrix

$$\mathcal{R}_{l,s,\phi} = \begin{bmatrix} \cos(\phi) & \frac{s}{l}\sin(\phi) \\ -\frac{l}{s}\sin(\phi) & \cos(\phi) \end{bmatrix}$$

is an elliptical rotation along the ellipse $\frac{x^2}{s^2} + \frac{y^2}{l^2} = d$ with $d > 0$.

Remark A.6. The definition is adopted from [61]. All similar ellipses have identical elliptical rotation matrices [61]. When $s = l$, then $\mathcal{R}_{l,s,\phi}$ simply becomes circular rotation.

Different from circular rotation which is isometry, elliptical rotation does not preserves angle and distance. Given any $x \in \mathbb{R}^n$ and its rotated point $x_+ = \mathcal{R}_{l,s,\phi}x$, the angle between x_+, x and the ratio $\frac{\|x_+\|}{\|x\|}$ depend on x. Given an elliptical rotation, let $e = (\frac{s^2}{l^2} - 1)/(\frac{s^2}{l^2} + 1)$ and $\zeta = \arccos(-e\cos(\phi))$.

Proposition A.7. *Let $\mathcal{R}_{l,s,\phi}$ be an elliptical rotation for some $\phi \in]0, \pi[$ and $l, s > 0$. Given an arbitrary point $x \neq 0$ and its rotated point $x_+ = \mathcal{R}_{l,s,\phi}x$, there holds*

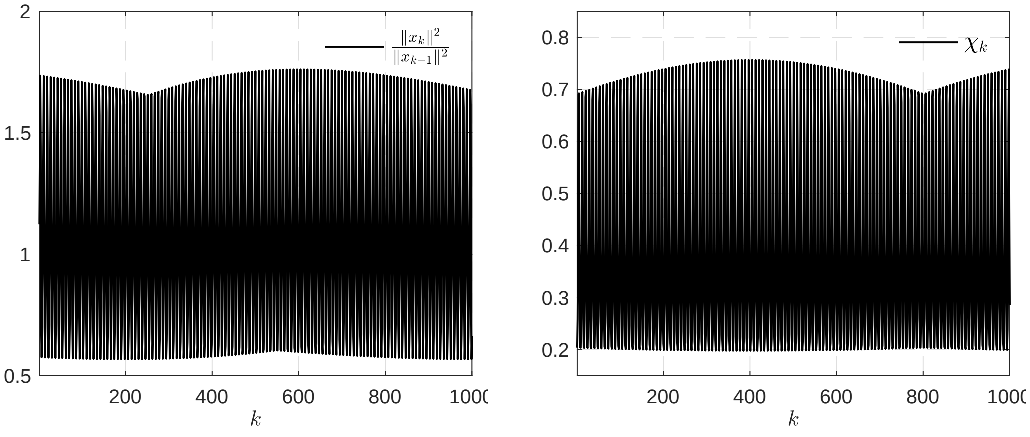
- *The ratio $\frac{\|x_+\|^2}{\|x\|^2} \in \left[\frac{e\cos(\zeta-\phi)+1}{e\cos(\zeta+\phi)+1}, \frac{e\cos(\zeta+\phi)+1}{e\cos(\zeta-\phi)+1}\right]$.*
- *Let χ be the angle between x and x_+ respect to 0, we have $\chi \in [\underline{\chi}, \overline{\chi}]$ with $\cos(\overline{\chi}) = \frac{a\cos(\phi)-b}{\sqrt{\sin^2(\phi)+(a\cos(\phi)-b)^2}}$ and $\cos(\underline{\chi}) = \frac{a\cos(\phi)+b}{\sqrt{\sin^2(\phi)+(a\cos(\phi)+b)^2}}$ where $a = \frac{s}{2l} + \frac{l}{2s}, b = |\frac{s}{2l} - \frac{l}{2s}|$.*

Remark A.8. When $s/l = 1$, then $\mathcal{R}_{l,s,\phi}$ becomes circular rotation, consequently we get $\frac{\|x_+\|^2}{\|x\|^2} = 1$ and $\chi = \phi$.

Example A.6. In this example, we consider an elliptical rotation parameterized by $l = 2, s = 1$ and $\phi = \frac{\pi}{30.01}$. Consider the sequence $\{x_k\}_{k\in\mathbb{N}}$ generated by the rotation $x_k = \mathcal{R}_{l,s,\phi}x_{k-1}$ with x_0 chosen arbitrarily, we study the ratio $\frac{\|x_k\|^2}{\|x_{k-1}\|^2}$ and the angle $\chi_k = \angle(x_k, x_{k-1})$

- We have $e = (\frac{s^2}{l^2} - 1)/(\frac{s^2}{l^2} + 1) = \frac{3}{5}$ and $\zeta = \arccos(-\frac{3\cos(\phi)}{5})$, consequently $\frac{\|x_k\|^2}{\|x_{k-1}\|^2} \in [0.5679, 1.7608]$.
- For the angle χ_k, we have that $\chi_k \in [0.1980, 0.7564]$.

The values of $\frac{\|x_k\|^2}{\|x_{k-1}\|^2}$ and χ_k along k are shown below in Figure 17.



(a) $\|x_k\|^2/\|x_{k-1}\|^2$

(b) $\cos(\chi_k)$

Figure 17: Range of $\frac{\|x_k\|^2}{\|x_{k-1}\|^2}$ and $\cos(\chi_k)$ for elliptical rotation.

Proof of Proposition A.7. Since $x \neq 0$, there exit $\beta \geq 0$ and $L, S > 0$ such that $S/L = s/l$ and $x = \begin{pmatrix} S\cos(\beta) \\ L\sin(\beta) \end{pmatrix}$. Then

$$x_+ = \mathfrak{R}_{l,s,\phi}x = \begin{bmatrix} \cos(\phi) & \frac{s}{l}\sin(\phi) \\ -\frac{l}{s}\sin(\phi) & \cos(\phi) \end{bmatrix} \begin{pmatrix} S\cos(\beta) \\ L\sin(\beta) \end{pmatrix} = \begin{pmatrix} S\cos(\phi)\cos(\beta) + S\sin(\phi)\sin(\beta) \\ L\cos(\phi)\sin(\beta) - L\sin(\phi)\cos(\beta) \end{pmatrix} = \begin{pmatrix} S\cos(\phi-\beta) \\ -L\sin(\phi-\beta) \end{pmatrix}. \tag{A.6}$$

We first prove the range of $\frac{\|x_+\|}{\|x\|}$,

$$q(\beta) = \frac{\|x_+\|^2}{\|x\|^2} = \frac{S^2\cos^2(\beta) + L^2\sin^2(\beta)}{S^2\cos^2(\phi-\beta) + L^2\sin^2(\phi-\beta)} = \frac{(\frac{s^2}{l^2}-1)\cos^2(\beta)+1}{(\frac{s^2}{l^2}-1)\cos^2(\phi-\beta)+1} = \frac{(\frac{s^2}{l^2}-1)\cos(2\beta)+\frac{s^2}{l^2}+1}{(\frac{s^2}{l^2}-1)\cos(2\phi-2\beta)+\frac{s^2}{l^2}+1}.$$

Denote $e = (\frac{s^2}{l^2}-1)/(\frac{s^2}{l^2}+1)$, then we get from above that $q(\beta) = \frac{e\cos(2\beta)+1}{e\cos(2\beta-2\phi)+1}$ whose derivative with respective to β reads

$$q'(\beta) = \frac{-2e^2\sin(2\phi) - 2e(\sin(2\beta) - \sin(2\beta-2\phi))}{(e\cos(2\beta-2\phi)+1)^2}.$$

Solving $q'(\beta) = 0$ we get

$$\begin{aligned} 0 = e\sin(2\phi) + \sin(2\beta) - \sin(2\beta-2\phi) \quad &\Longleftrightarrow \quad 0 = 2\sin(\phi)\big(e\cos(\phi) + \cos(2\beta-\phi)\big) \\ &\Longleftrightarrow \quad -e\cos(\phi) = \cos(2\beta-\phi). \end{aligned}$$

Denote $\zeta = \arccos(-e\cos(\phi))$, then the choices of β such that $q'(\beta) = 0$ hence $q(\beta)$ reaches extreme values are

$$\beta_{\max} = \frac{\zeta+\phi}{2} \quad \text{and} \quad \beta_{\min} = \pi - \frac{\zeta-\phi}{2}.$$

Consequently we get

$$\max_{\beta\in[0,2\pi]} q(\beta) = q(\beta_{\max}) = \frac{e\cos(\zeta+\phi)+1}{e\cos(\zeta-\phi)+1} \quad \text{and} \quad \min_{\beta\in[0,2\pi]} q(\beta) = q(\beta_{\min}) = \frac{e\cos(\zeta-\phi)+1}{e\cos(\zeta+\phi)+1},$$

which is the range of $\frac{\|x_+\|^2}{\|x\|^2}$.

For the angle χ between x_+ and x, from (A.6) we get the following inner product

$$\langle x_+, x\rangle = S^2\cos(\beta)\cos(\phi-\beta) - L^2\sin(\phi-\beta)\sin(\beta) = \frac{S^2+L^2}{2}\cos(\phi) + \frac{S^2-L^2}{2}\cos(\phi-2\beta), \tag{A.7}$$

which means

$$\cos(\chi) = \frac{\langle x_+, x\rangle}{\|x_+\|\|x\|} = \frac{\frac{s/l+l/s}{2}\cos(\phi) + \frac{s/l-l/s}{2}\cos(\phi-2\beta)}{\sqrt{\sin^2(\phi-\beta)+s^2/l^2\cos^2(\phi-\beta)}\sqrt{l^2/s^2\sin^2(\beta)+\cos^2(\beta)}}.$$

Since ϕ and s/l are constant, consider the function of β:

$$\begin{aligned}
\ell(\beta) &\overset{\text{def}}{=} \big(\sin^2(\phi-\beta)+\tfrac{s^2}{l^2}\cos^2(\phi-\beta)\big)\big(\tfrac{l^2}{s^2}\sin^2(\beta)+\cos^2(\beta)\big)\\
&= \tfrac{l^2}{s^2}\sin^2(\beta)\sin^2(\phi-\beta)+\cos^2(\beta)\sin^2(\phi-\beta)+\sin^2(\beta)\cos^2(\phi-\beta)+\tfrac{s^2}{l^2}\cos^2(\beta)\cos^2(\phi-\beta)\\
&= \big(\sin(\beta)\cos(\phi-\beta)+\cos(\beta)\sin(\phi-\beta)\big)^2-2\sin(\beta)\cos(\phi-\beta)\cos(\beta)\sin(\phi-\beta)\\
&\qquad+\tfrac{l^2}{s^2}\sin^2(\beta)\sin^2(\phi-\beta)+\tfrac{s^2}{l^2}\cos^2(\beta)\cos^2(\phi-\beta)\\
&= \big(\sin(\beta)\cos(\phi-\beta)+\cos(\beta)\sin(\phi-\beta)\big)^2+\big(\tfrac{s}{l}\cos(\beta)\cos(\phi-\beta)-\tfrac{l}{s}\sin(\beta)\sin(\phi-\beta)\big)^2\\
&= \sin^2(\phi)+\big(\tfrac{s}{l}\cos(\beta)\cos(\phi-\beta)-\tfrac{s}{l}\sin(\beta)\sin(\phi-\beta)-(\tfrac{l}{s}-\tfrac{s}{l})\sin(\beta)\sin(\phi-\beta)\big)^2\\
&= \sin^2(\phi)+\big(\tfrac{s}{l}\cos(\phi)-(\tfrac{l}{s}-\tfrac{s}{l})\sin(\beta)\sin(\phi-\beta)\big)^2\\
&= \sin^2(\phi)+\Big(\tfrac{s}{l}\cos(\phi)-(\tfrac{l}{s}-\tfrac{s}{l})\frac{\cos(\phi-2\beta)-\cos(\phi)}{2}\Big)^2\\
&= \sin^2(\phi)+\Big(\tfrac{s}{l}\cos(\phi)+(\tfrac{l}{s}-\tfrac{s}{l})\frac{\cos(\phi)}{2}-(\tfrac{l}{s}-\tfrac{s}{l})\frac{\cos(\phi-2\beta)}{2}\Big)^2\\
&= \sin^2(\phi)+\big((\tfrac{l}{2s}+\tfrac{s}{2l})\cos(\phi)-(\tfrac{l}{2s}-\tfrac{s}{2l})\cos(\phi-2\beta)\big)^2.
\end{aligned}$$

Therefore, we have

$$\cos(\chi)=\frac{(\frac{s}{2l}+\frac{l}{2s})\cos(\phi)+(\frac{s}{2l}-\frac{l}{2s})\cos(\phi-2\beta)}{\sqrt{\sin^2(\phi)+((\frac{l}{2s}+\frac{s}{2l})\cos(\phi)+(\frac{s}{2l}-\frac{l}{2s})\cos(\phi-2\beta))^2}}.$$

Consider the following function

$$f(x)=\frac{x}{\sqrt{\sin^2(\phi)+x^2}}.$$

It is easy to verify that the derivative $f'(x)=\frac{\sin^2(\phi)}{(\sin^2(\phi)+x^2)^{3/2}}>0$ holds for all $x\in\mathbb{R}$, hence $f(x)$ is monotonically increasing. Now let $x=(\frac{s}{2l}+\frac{l}{2s})\cos(\phi)+(\frac{s}{2l}-\frac{l}{2s})\cos(\phi-2\beta)$, we have

$$x\in\big[\big(\tfrac{s}{2l}+\tfrac{l}{2s}\big)\cos(\phi)-|\tfrac{s}{2l}-\tfrac{l}{2s}|,\big(\tfrac{s}{2l}+\tfrac{l}{2s}\big)\cos(\phi)+|\tfrac{s}{2l}-\tfrac{l}{2s}|\big],$$

which further implies

$$\frac{(\frac{s}{2l}+\frac{l}{2s})\cos(\phi)-|\frac{s}{2l}-\frac{l}{2s}|}{\sqrt{\sin^2(\phi)+((\frac{s}{2l}+\frac{l}{2s})\cos(\phi)-|\frac{s}{2l}-\frac{l}{2s}|)^2}}\le\cos(\chi)\le\frac{(\frac{s}{2l}+\frac{l}{2s})\cos(\phi)+|\frac{s}{2l}-\frac{l}{2s}|}{\sqrt{\sin^2(\phi)+((\frac{s}{2l}+\frac{l}{2s})\cos(\phi)+|\frac{s}{2l}-\frac{l}{2s}|)^2}}.$$

Since $\cos(\chi)$ is monotonic decreasing in $[0,\pi/2]$, we obtain the claimed result. □

Given an angle $\psi\in]0,\pi]$, define the circular rotation $\mathcal{R}_\psi=\begin{bmatrix}\cos(\psi) & -\sin(\psi)\\ \sin(\psi) & \cos(\psi)\end{bmatrix}$. For the composite rotation $\mathcal{R}_\psi\mathcal{R}_{l,s,\phi}$, we have the following.

Proposition A.9. *Let $\mathcal{R}_{l,s,\phi}$ be an elliptical rotation for $\phi\in]0,\pi/2]$ and $l,s>0$ and $\mathcal{R}_\psi$ be a circular rotation such that*

$$\big((\tfrac{l}{s}+\tfrac{s}{l})\sin(\psi)\sin(\phi)+2\cos(\psi)\cos(\phi)\big)^2-4<0.$$

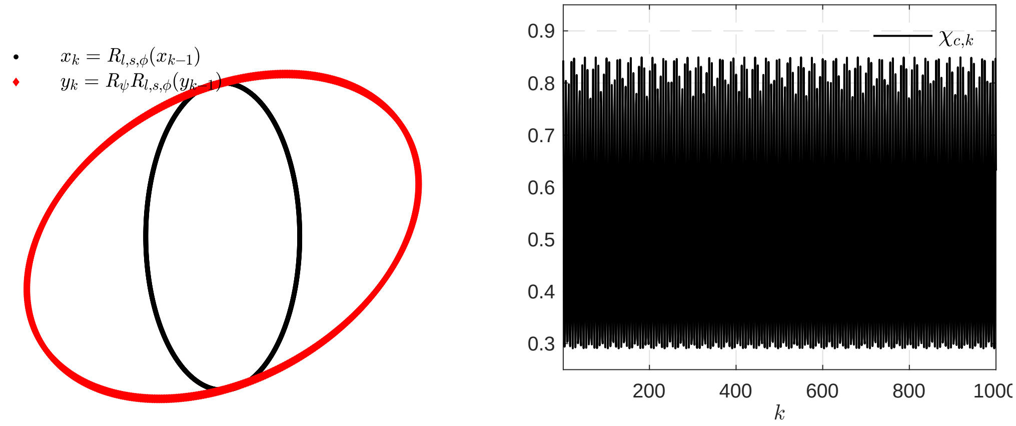
Given an arbitrary point $x\neq 0$ and its rotated point $x_+=\mathcal{R}_\psi\mathcal{R}_{l,s,\phi}x$,

- *The ratio $\frac{\|x_+\|^2}{\|x\|^2}\in\big[\frac{e\cos(\zeta-\phi)+1}{e\cos(\zeta+\phi)+1},\frac{e\cos(\zeta+\phi)+1}{e\cos(\zeta-\phi)+1}\big]$ as in Proposition A.7.*
- *Let $\chi_c=\angle(x,x_+)$ be the angle between x and x_+, then $\chi_c\in[\psi-\overline{\chi},\psi-\underline{\chi}]$ where $\underline{\chi},\overline{\chi}$ are as defined in Proposition A.7.*

Example A.6 (Continued). We continue Example A.6 by compositing $\mathcal{R}_{l,s,\phi}$ with a circular rotation. Let $\psi=\frac{\pi}{3}$ and consider the sequence $\{y_k\}_{k\in\mathbb{N}}$ generated by the rotation $y_k=\mathcal{R}_\psi\mathcal{R}_{l,s,\phi}y_{k-1}$ with $y_0=x_0$, we consider the trajectory of the sequence $\{y_k\}_{k\in\mathbb{N}}$ and angle $\chi_{c,k}=\angle(y_k,y_{k-1}),\vartheta_{c,k}=\angle(y_k,y_k-y_{k-1})$:

- For the elliptical rotation and the composite rotation, trajectories of the sequences $\{x_k\}_{k\in\mathbb{N}},\{y_k\}_{k\in\mathbb{N}}$ are shown below in Figure 18 (a). Note that the trajectory of $\{y_k\}_{k\in\mathbb{N}}$ is not equal to rotating the that of $\{x_k\}_{k\in\mathbb{N}}$ using $\mathcal{R}_\psi$.

- For the angle $\chi_{c,k}$, we have that $\chi_{c,k} \in [\psi - \overline{\chi}, \psi - \underline{\chi}] = [0.2908, 0.8492]$.



(a) Trajectory of x_k (b) $\cos(\chi_{c,k})$

Figure 18: Trajectories of sequences $\{x_k\}_{k\in\mathbb{N}}, \{y_k\}_{k\in\mathbb{N}}$, and the range of $\chi_{c,k}$.

Proof. Denote $\mathcal{R} = \mathcal{R}_\psi \mathcal{R}_{l,s,\phi}$, then we have

$$\mathcal{R} = \begin{bmatrix} \cos(\psi)\cos(\phi) + \frac{l}{s}\sin(\psi)\sin(\phi) & \frac{s}{l}\cos(\psi)\sin(\phi) - \sin(\psi)\cos(\phi) \\ \sin(\psi)\cos(\phi) - \frac{l}{s}\cos(\psi)\sin(\phi) & \frac{s}{l}\sin(\psi)\sin(\phi) + \cos(\psi)\cos(\phi) \end{bmatrix} = \begin{bmatrix} \mathcal{R}_{1,1} & \mathcal{R}_{1,2} \\ \mathcal{R}_{2,1} & \mathcal{R}_{2,2} \end{bmatrix}$$

and that

$$\mathcal{R}_{1,1} - \mathcal{R}_{2,2} = \left(\tfrac{l}{s} - \tfrac{s}{l}\right)\sin(\psi)\sin(\phi),$$
$$\mathcal{R}_{1,2}\mathcal{R}_{2,1} = \left(\tfrac{s}{l} + \tfrac{l}{s}\right)\cos(\psi)\sin(\phi)\sin(\psi)\cos(\phi) - \cos^2(\psi)\sin^2(\phi) - \sin^2(\psi)\cos^2(\phi).$$

The characteristic polynomial of $\mathcal{R}$ reads

$$x^2 - (\mathcal{R}_{1,1} + \mathcal{R}_{2,2})x + \mathcal{R}_{1,1}\mathcal{R}_{2,2} - \mathcal{R}_{1,2}\mathcal{R}_{2,1} = 0,$$

whose discriminant is

$$\begin{aligned} \Delta &= (\mathcal{R}_{1,1} + \mathcal{R}_{2,2})^2 - 4(\mathcal{R}_{1,1}\mathcal{R}_{2,2} - \mathcal{R}_{1,2}\mathcal{R}_{2,1}) \\ &= (\mathcal{R}_{1,1} - \mathcal{R}_{2,2})^2 + 4\mathcal{R}_{1,2}\mathcal{R}_{2,1} \\ &= \left(\tfrac{l}{s} - \tfrac{s}{l}\right)^2 \sin^2(\psi)\sin^2(\phi) + 4\left(\tfrac{s}{l} + \tfrac{l}{s}\right)\cos(\psi)\sin(\phi)\sin(\psi)\cos(\phi) - 4\cos^2(\psi)\sin^2(\phi) - 4\sin^2(\psi)\cos^2(\phi) \\ &= \Big(\left(\tfrac{l}{s} + \tfrac{s}{l}\right)\sin(\psi)\sin(\phi) + 2\cos(\psi)\cos(\phi)\Big)^2 - 4. \end{aligned}$$

When $\Delta < 0$, $\mathcal{R}$ admits two complex eigenvalues, meaning that $\mathcal{R}$ is a rotation.

Let $y = \mathcal{R}_{l,s,\phi}x$ and let β be the angle between x, y, then owing to Proposition A.7, we have $\beta \in [\underline{\chi}, \overline{\chi}]$. The angle between x_+ and y is ψ which is straightforward owing to $\mathcal{R}_\psi$. Since $\mathcal{R}_{l,s,\phi}$ is clockwise rotation and $\mathcal{R}_\psi$ is counterclockwise, the claimed result follows immediately. □

A.3.2 Type III linear system

Let $m_1, m_2 \in \mathbb{N}_+$ such that $m_1 + m_2 = n$, let $A \in \mathbb{R}^{m_1\times m_1}, B \in \mathbb{R}^{m_2\times m_2}$ be symmetric and $C \in \mathbb{R}^{m_2\times m_1}$. Define the following 2×2 block matrix

$$M \overset{\text{def}}{=} \begin{bmatrix} A & -\delta C^T \\ \tau C & B \end{bmatrix}. \tag{A.8}$$

Definition A.7 (Type III matrix). $M \in \mathbb{R}^{n\times n}$ is a 2×2 block matrix defined by (A.8), with all its eigenvalues lying in the complex unit disc.

For the sake of brevity, we assume henceforth that $m_1 = m_2 = m = \frac{n}{2}$[5]. Denote $S_A = (a_i)_{i=1,\dots,m}, S_B = (b_i)_{i=1,\dots,m}$ and $S_C = (c_i)_{i=1,\dots,m} \in \mathbb{R}^m$ the singular values of A, B and C in *descending* order, respectively. For each $i = 1, \dots, m$, define the 2×2 matrix D_i by $D_i = \begin{bmatrix} a_i & -\delta c_i \\ \tau c_i & b_i \end{bmatrix}$. It is trivial to show that the eigenvalues of D_i read $\frac{1}{2}\big((a_i + b_i) \pm \sqrt{(a_i - b_i)^2 - 4\delta\tau c_i^2}\big)$. We impose the following assumptions.

Assumption A.10. *Let $C = Y\mathrm{diag}(S_C)X^T$ be the SVD of C, suppose that A and B can be diagonalized by X and Y respectively, that is there holds $\mathrm{diag}(S_A) = X^TAX$ and $\mathrm{diag}(S_B) = Y^TBY$. For each $i = 1, \dots, m$:*

(i) *If the eigenvalues are real, i.e. $(a_i - b_i)^2 - 4\delta\tau c_i^2 \geq 0$, then they are either 0 or 1;*

(ii) *If $(a_i - b_i)^2 - 4\delta\tau c_i^2 < 0$, then $\delta\tau c_i^2 + a_ib_i < 1$. Moreover, there exists $1 \leq q \leq d \leq m$ such that $D_i = D_j, 1 \leq i, j \leq q$ and $\delta\tau c_1^2 + a_1b_1 = \cdots = \delta\tau c_q^2 + a_qb_q > \delta\tau c_{q+1}^2 + a_{q+1}b_{q+1} \geq \cdots \geq \delta\tau c_d^2 + a_db_d > 0$.*

Let $\sigma = \sqrt{\delta\tau c_1^2 + a_1b_1}$ and $\eta = \frac{\sqrt{\delta\tau c_{q+1}^2 + a_{q+1}b_{q+1}}}{\sigma}$, we have the following proposition. Again, let $M^\infty \overset{\text{def}}{=} \lim_{k\to+\infty} M^k$.

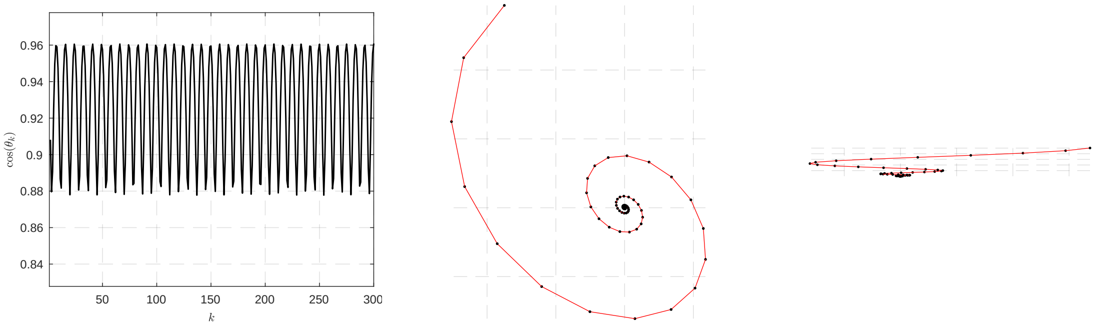
Proposition A.11. *Consider the linear system (A.2) whose M is a Type III matrix define in Definition A.7, suppose that Assumption A.10 holds. Then*

(i) *M^∞ is a symmetric matrix with eigenvalues being either 0 or 1, and $v_0 \in \ker(M)$.*

(ii) *$\theta_k \in \big[\psi - \overline{\alpha}, \psi - \underline{\alpha}\big]$, where $\cos(\overline{\chi}) = \frac{(\frac{s}{2l} + \frac{l}{2s})\cos(\phi) - |\frac{s}{2l} - \frac{l}{2s}|}{\sqrt{\sin^2(\phi) + ((\frac{s}{2l} + \frac{l}{2s})\cos(\phi) - |\frac{s}{2l} - \frac{l}{2s}|)^2}}$ and $\cos(\underline{\chi}) = \frac{(\frac{s}{2l} + \frac{l}{2s})\cos(\phi) + |\frac{s}{2l} - \frac{l}{2s}|}{\sqrt{\sin^2(\phi) + ((\frac{s}{2l} + \frac{l}{2s})\cos(\phi) + |\frac{s}{2l} - \frac{l}{2s}|)^2}}$, and $\psi = \mathrm{arccot}\big(\frac{(\tau-\delta)c_1}{b_1 - a_1}\big)$, $\phi = \arccos\big(\frac{(\delta+\tau)c_1\sin(\psi) + (a_1+b_1)\cos(\psi)}{2\sigma}\big)$ and $\frac{s}{l} = \frac{b_1}{\sigma\sin(\psi)\sin(\phi)} - \cot(\psi)\cot(\phi)$.*

Remark A.12. When $\delta = \tau = 1$, then we have $\psi = \frac{\pi}{2}, \phi = \arccos\big(\frac{c_1}{\sqrt{c_1^2 + a_1b_1}}\big)$ and $\frac{s}{l} = \frac{b_1}{\sin(\phi)\sqrt{c_1^2 + a_1b_1}}$.

Example A.8. Let $\delta = \tau = 1$ and $a, b, c, d > 0$, and let $M = U \begin{bmatrix} a & -c & \\ c & b & \\ & & d \end{bmatrix} U^T$ where U is an orthogonal matrix in $\mathbb{R}^{3\times 3}$. M has three eigenvalues: $\frac{(a+b)+\sqrt{(a-b)^2-4c^2}}{2}, \frac{(a+b)-\sqrt{(a-b)^2-4c^2}}{2}$ and d. Note that the magnitude of both complex eigenvalues is $\sqrt{ab + c^2}$. We consider the following choice of a, b, c, d: $(a, b, c) = (0.95, 0.75, 0.35)$ and $d = 0.99\sqrt{ab + c^2}$. The observations are shown in Figure 19, where the first figure shows the oscillation behavior of $\cos\theta_k$, and the trajectories of $\{x_k\}_{k\in\mathbb{N}}$ from different perspectives are provided in the 2nd and 3rd figures.



(a) $\cos(\theta_k)$ (b) Top view (c) Side view

Figure 19: Oscillation of $\cos(\theta_k)$ and trajectory of $\{x_k\}_{k\in\mathbb{N}}$. (a): Oscillation of $\cos(\theta_k)$; (b) Top view of trajectory of $\{x_k\}_{k\in\mathbb{N}}$; (c) Side view of trajectory of $\{x_k\}_{k\in\mathbb{N}}$.

Proof of Proposition A.11. Owing to Assumption A.10, denote $\Sigma_A = \mathrm{diag}(S_A), \Sigma_B = \mathrm{diag}(S_B)$ and $\Sigma_C = \mathrm{diag}(S_C)$, then we have for M that

$$M = \begin{bmatrix} A & -C^T \\ C & B \end{bmatrix} = \begin{bmatrix} X\Sigma_AX^T & -\delta X\Sigma_CY^T \\ \tau Y\Sigma_CX^T & Y\Sigma_BY^T \end{bmatrix} = \begin{bmatrix} X & \\ & Y \end{bmatrix} \begin{bmatrix} \Sigma_A & -\delta\Sigma_C \\ \tau\Sigma_C & \Sigma_B \end{bmatrix} \begin{bmatrix} X^T & \\ & Y^T \end{bmatrix}. \tag{A.9}$$

[5] For the case of $m_1 \neq m_2$ or n is odd, we can apply the zero padding trick.

Define the following matrix

$$\Sigma \overset{\text{def}}{=} \begin{bmatrix} \Sigma_A & -\delta\Sigma_C \\ \tau\Sigma_C & \Sigma_B \end{bmatrix}$$

which is block diagonal matrix. For each $i = 1, ..., p$, define the 2×2 matrix $D_i = \begin{bmatrix} a_i & -\delta c_i \\ \tau c_i & b_i \end{bmatrix}$. Owing to Assumption A.10, when $(a_i - b_i)^2 - 4\delta\tau c_i^2 < 0$, the eigenvalues of D_i, *i.e.* $\frac{(a_i+b_i)\pm\sqrt{(a_i-b_i)^2-4\delta\tau c_i^2}}{2}$, are complex and their magnitudes is $\delta\tau c_i^2 + a_i b_i < 1$. Therefore, we have $\lim_{k\to+\infty} D_i^k = 0$ owing to spectral theorem. This further implies

$$M^\infty \overset{\text{def}}{=} \lim_{k\to+\infty} M^k = \begin{bmatrix} Y & \\ & X \end{bmatrix} \begin{bmatrix} \mathrm{Id}_r & \\ & 0_{n-r} \end{bmatrix} \begin{bmatrix} Y^T & \\ & X^T \end{bmatrix},$$

where r is the multiplicities of eigenvalue 1.

Following the arguments of the proof of Proposition A.4, we have that $v_k \in \ker(M^\infty)$ for all $k \in \mathbb{N}$. Again, define $\widetilde{M} \overset{\text{def}}{=} M - M^\infty$. Based on (A.9) and block-diagonal nature of Σ, there exists an elementary transformation matrix Z such that, let $d = m - r$

$$\widetilde{M} = \begin{bmatrix} X & \\ & Y \end{bmatrix} Z \begin{bmatrix} D_1 & & & \\ & \ddots & & \\ & & D_d & \\ & & & 0_{2r} \end{bmatrix} Z^T \begin{bmatrix} X^T & \\ & Y^T \end{bmatrix} = W\Gamma W^T, \tag{A.10}$$

where $W = \begin{bmatrix} X & \\ & Y \end{bmatrix} Z$ and Γ is the block diagonal matrix. The order of $D_i, i = 1, ..., d$ is such that it complies with Assumption A.10. Consider the following decomposition of Γ

$$\Gamma_1 = \begin{bmatrix} D_1 & & & \\ & \ddots & & \\ & & D_q & \\ & & & 0_{n-2q} \end{bmatrix} \quad \text{and} \quad \Gamma_2 = \Gamma - \Gamma_1.$$

Let $\sigma = \sqrt{\delta\tau c_1^2 + a_1 b_1}$ and $\eta = \frac{\sqrt{\delta\tau c_{q+1}^2 + a_{q+1} b_{q+1}}}{\sigma}$, then $\frac{1}{\sigma^k}\Gamma_2^k = O(\eta^k) \to 0$.

Follow the proof of Proposition A.4, θ_k eventually is determined by the rotation property of Γ_1. Clearly, there exist some $\psi, \phi \in [0, \pi/2]$ and $l, s > 0$ such that

$$\frac{1}{\sigma} D_1 = \frac{1}{\sigma} \begin{bmatrix} a_1 & -\delta c_1 \\ \tau c_1 & b_1 \end{bmatrix} = \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix} \begin{bmatrix} \cos(\phi) & \frac{s}{l}\sin(\phi) \\ -\frac{l}{s}\sin(\phi) & \cos(\phi) \end{bmatrix}.$$

Consequently, we get

$$\begin{aligned} \cos(\psi)\cos(\phi) + \tfrac{l}{s}\sin(\psi)\sin(\phi) &= \frac{a_1}{\sigma}, \\ \tfrac{s}{l}\cos(\psi)\sin(\phi) - \sin(\psi)\cos(\phi) &= -\frac{\delta c_1}{\sigma}, \\ \sin(\psi)\cos(\phi) - \tfrac{l}{s}\cos(\psi)\sin(\phi) &= \frac{\tau c_1}{\sigma}, \\ \tfrac{s}{l}\sin(\psi)\sin(\phi) + \cos(\psi)\cos(\phi) &= \frac{b_1}{\sigma}, \end{aligned} \tag{A.11}$$

which yields

$$\psi = \operatorname{arccot}\left(\tfrac{(\tau-\delta)c_1}{b_1-a_1}\right), \ \phi = \arccos\left(\tfrac{(\delta+\tau)c_1\sin(\psi)+(a_1+b_1)\cos(\psi)}{2\sigma}\right) \quad \text{and} \quad \tfrac{s}{l} = \tfrac{b_1}{\sin(\psi)\sin(\phi)\sigma} - \cot(\psi)\cot(\phi).$$

This means that $\frac{1}{\sigma}D_1$ is the composite rotation of Proposition A.9, and so is $\frac{1}{\sigma}\Gamma_1$. Therefore, invoke the result of Proposition A.7 and A.9, we have $\theta_k \in \left[\psi - \overline{\chi}, \psi - \underline{\chi}\right]$ with

$$\cos(\overline{\chi}) = \frac{(\frac{s}{2l} + \frac{l}{2s})\cos(\phi) - |\frac{s}{2l} - \frac{l}{2s}|}{\sqrt{\sin^2(\phi) + ((\frac{s}{2l} + \frac{l}{2s})\cos(\phi) - |\frac{s}{2l} - \frac{l}{2s}|)^2}} \quad \text{and} \quad \cos(\underline{\chi}) = \frac{(\frac{s}{2l} + \frac{l}{2s})\cos(\phi) + |\frac{s}{2l} - \frac{l}{2s}|}{\sqrt{\sin^2(\phi) + ((\frac{s}{2l} + \frac{l}{2s})\cos(\phi) + |\frac{s}{2l} - \frac{l}{2s}|)^2}}. \quad \square$$

B Proofs of Section 3

B.1 Riemannian Geometry

Let $\mathcal{M}$ be a C^2-smooth embedded submanifold of $\mathbb{R}^n$ around a point x. With some abuse of terminology, we shall state C^2-manifold instead of C^2-smooth embedded submanifold of $\mathbb{R}^n$. The natural embedding of a submanifold $\mathcal{M}$ into $\mathbb{R}^n$ permits to define a Riemannian structure and to introduce geodesics on $\mathcal{M}$, and we simply say $\mathcal{M}$ is a Riemannian manifold. We denote respectively $\mathcal{T}_{\mathcal{M}}(x)$ and $\mathcal{N}_{\mathcal{M}}(x)$ the tangent and normal space of $\mathcal{M}$ at point near x in $\mathcal{M}$.

Exponential map Geodesics generalize the concept of straight lines in $\mathbb{R}^n$, preserving the zero acceleration characteristic, to manifolds. Roughly speaking, a geodesic is locally the shortest path between two points on $\mathcal{M}$. We denote by $\mathfrak{g}(t;x,h)$ the value at $t \in \mathbb{R}$ of the geodesic starting at $\mathfrak{g}(0;x,h) = x \in \mathcal{M}$ with velocity $\dot{\mathfrak{g}}(t;x,h) = \frac{d\mathfrak{g}}{dt}(t;x,h) = h \in \mathcal{T}_{\mathcal{M}}(x)$ (which is uniquely defined). For every $h \in \mathcal{T}_{\mathcal{M}}(x)$, there exists an interval I around 0 and a unique geodesic $\mathfrak{g}(t;x,h) : I \to \mathcal{M}$ such that $\mathfrak{g}(0;x,h) = x$ and $\dot{\mathfrak{g}}(0;x,h) = h$. The mapping

$$\mathrm{Exp}_x : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{M},\ h \mapsto \mathrm{Exp}_x(h) = \mathfrak{g}(1;x,h),$$

is called *Exponential map*. Given $x, x' \in \mathcal{M}$, the direction $h \in \mathcal{T}_{\mathcal{M}}(x)$ we are interested in is the one such that $\mathrm{Exp}_x(h) = x' = \mathfrak{g}(1;x,h)$.

Parallel translation Given two points $x, x' \in \mathcal{M}$, let $\mathcal{T}_{\mathcal{M}}(x), \mathcal{T}_{\mathcal{M}}(x')$ be their corresponding tangent spaces. Define

$$\tau : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x'),$$

the parallel translation along the unique geodesic joining x to x', which is isomorphism and isometry with respect to the Riemannian metric.

Riemannian gradient and Hessian For a vector $v \in \mathcal{N}_{\mathcal{M}}(x)$, the Weingarten map of $\mathcal{M}$ at x is the operator $\mathfrak{W}_x(\cdot, v) : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x)$ defined by

$$\mathfrak{W}_x(\cdot, v) = -\mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)} \mathrm{d}V[h],$$

where V is any local extension of v to a normal vector field on $\mathcal{M}$. The definition is independent of the choice of the extension V, and $\mathfrak{W}_x(\cdot, v)$ is a symmetric linear operator which is closely tied to the second fundamental form of $\mathcal{M}$, see [26, Proposition II.2.1].

Let G be a real-valued function which is C^2 along the $\mathcal{M}$ around x. The covariant gradient of G at $x' \in \mathcal{M}$ is the vector $\nabla_{\mathcal{M}} G(x') \in \mathcal{T}_{\mathcal{M}}(x')$ defined by

$$\langle \nabla_{\mathcal{M}} G(x'), h \rangle = \frac{d}{dt} G\big(\mathcal{P}_{\mathcal{M}}(x' + th)\big)\big|_{t=0},\ \forall h \in \mathcal{T}_{\mathcal{M}}(x'),$$

where $\mathcal{P}_{\mathcal{M}}$ is the projection operator onto $\mathcal{M}$. The covariant Hessian of G at x' is the symmetric linear mapping $\nabla^2_{\mathcal{M}} G(x')$ from $\mathcal{T}_{\mathcal{M}}(x')$ to itself which is defined as

$$\langle \nabla^2_{\mathcal{M}} G(x') h, h \rangle = \frac{d^2}{dt^2} G\big(\mathcal{P}_{\mathcal{M}}(x' + th)\big)\big|_{t=0},\ \forall h \in \mathcal{T}_{\mathcal{M}}(x'). \tag{B.1}$$

This definition agrees with the usual definition using geodesics or connections [56]. Now assume that $\mathcal{M}$ is a Riemannian embedded submanifold of $\mathbb{R}^n$, and that a function G has a C^2-smooth restriction on $\mathcal{M}$. This can be characterized by the existence of a C^2-smooth extension (representative) of G, *i.e.* a C^2-smooth function $\widetilde{G}$ on $\mathbb{R}^n$ such that $\widetilde{G}$ agrees with G on $\mathcal{M}$. Thus, the Riemannian gradient $\nabla_{\mathcal{M}} G(x')$ is also given by

$$\nabla_{\mathcal{M}} G(x') = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla \widetilde{G}(x'), \tag{B.2}$$

and $\forall h \in \mathcal{T}_{\mathcal{M}}(x')$, the Riemannian Hessian reads

$$\begin{aligned} \nabla^2_{\mathcal{M}} G(x') h &= \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \mathrm{d}(\nabla_{\mathcal{M}} G)(x')[h] = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \mathrm{d}\big(x' \mapsto \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla_{\mathcal{M}} \widetilde{G}\big)[h] \\ &= \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla^2 \widetilde{G}(x') h + \mathfrak{W}_{x'}\big(h, \mathcal{P}_{\mathcal{N}_{\mathcal{M}}(x')} \nabla \widetilde{G}(x')\big), \end{aligned} \tag{B.3}$$

where the last equality comes from [1, Theorem 1]. When $\mathcal{M}$ is an affine or linear subspace of $\mathbb{R}^n$, then obviously $\mathcal{M} = x + \mathcal{T}_{\mathcal{M}}(x)$, and $\mathfrak{W}_{x'}(h, \mathcal{P}_{\mathcal{N}_{\mathcal{M}}(x')} \nabla \widetilde{G}(x')) = 0$, hence (B.3) reduces to

$$\nabla^2_{\mathcal{M}} G(x') = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')} \nabla^2 \widetilde{G}(x') \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x')}.$$

See [43, 26] for more materials on differential and Riemannian manifolds.

Below present the expressions of the Riemannian gradient and Hessian for the case of partly smooth functions relative to a C^2-smooth manifold.

Lemma B.1 (Riemannian gradient and Hessian). *If $R \in \mathrm{PSF}_x(\mathcal{M}_x)$, then for any point $x' \in \mathcal{M}_x$ near x*

$$\nabla_{\mathcal{M}_x} R(x') = \mathcal{P}_{T_{x'}}(\partial R(x')),$$

and this does not depend on the smooth representation of R on $\mathcal{M}_x$. In turn, for all $h \in T_{x'}$, let $\widetilde{R}$ be a smooth representative of R on $\mathcal{M}_x$,

$$\nabla^2_{\mathcal{M}_x} R(x')h = \mathcal{P}_{T_{x'}} \nabla^2 \widetilde{R}(x')h + \mathfrak{W}_{x'}\big(h, \mathcal{P}_{T_{x'}^{\perp}} \nabla \widetilde{R}(x')\big),$$

where $\mathfrak{W}_x(\cdot,\cdot) : T_x \times T_x^{\perp} \to T_x$ is the Weingarten map of $\mathcal{M}_x$ at x.

The result of Lemma B.1 implies that we can linearize the proximity operators along the C^2-smooth manifold, which is discussed in Lemma B.6.

Lemma B.2 ([46, Lemma 5.1]). *Let $\mathcal{M}$ be a C^2-smooth manifold around x. Then for any $x' \in \mathcal{M} \cap \mathcal{N}$, where $\mathcal{N}$ is a neighborhood of x, the projection operator $\mathcal{P}_{\mathcal{M}}(x')$ is uniquely valued and C^1 around x, and thus*

$$x' - x = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x' - x) + o(\|x' - x\|).$$

If moreover $\mathcal{M} = x + \mathcal{T}_{\mathcal{M}}(x)$ is an affine subspace, then $x' - x = \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x' - x)$.

Lemma B.3 ([47, Lemma B.1]). *Let $x \in \mathcal{M}$, and x_k a sequence converging to x in $\mathcal{M}$. Denote $\tau_k : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x_k)$ be the parallel translation along the unique geodesic joining x to x_k. Then, for any bounded vector $u \in \mathbb{R}^n$, we have*

$$(\tau_k^{-1} \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x_k)} - \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)})u = o(\|u\|).$$

Lemma B.4 ([47, Lemma B.2]). *Let x, x' be two close points in $\mathcal{M}$, denote $\tau : \mathcal{T}_{\mathcal{M}}(x) \to \mathcal{T}_{\mathcal{M}}(x')$ the parallel translation along the unique geodesic joining x to x'. The Riemannian Taylor expansion of $\Phi \in C^2(\mathcal{M})$ around x reads,*

$$\tau^{-1} \nabla_{\mathcal{M}} \Phi(x') = \nabla_{\mathcal{M}} \Phi(x) + \nabla^2_{\mathcal{M}} \Phi(x) \mathcal{P}_{\mathcal{T}_{\mathcal{M}}(x)}(x' - x) + o(\|x' - x\|).$$

B.2 Linearization of proximal mapping

When dealing with non-smooth optimization, one fundamental result, provided by partial smoothness, is the linearization of proximal mapping. We first discuss the property of the Riemannian Hessian of a partly smooth function. Let $R \in \Gamma_0(\mathbb{R}^n)$ be partly smooth at $\bar{x}$ relative to $\mathcal{M}_{\bar{x}}$ and $\bar{u} \in \partial R(\bar{x})$, define the following smooth perturbation of R

$$\overline{R}(x) \overset{\text{def}}{=} R(x) - \langle x, \bar{u} \rangle, \tag{B.4}$$

whose Riemannian Hessian at $\bar{x}$ reads $H_{\overline{R}} \overset{\text{def}}{=} \mathcal{P}_{T_{\bar{x}}} \nabla^2_{\mathcal{M}_{\bar{x}}} \overline{R}(\bar{x}) \mathcal{P}_{T_{\bar{x}}}$.

Lemma B.5 ([47, Lemma 4.2]). *Let $R \in \Gamma_0(\mathbb{R}^n)$ be partly smooth at $\bar{x}$ relative to $\mathcal{M}_{\bar{x}}$, then $H_{\overline{R}}$ is symmetric positive semi-definite if either of the following is true:*

- *$\bar{u} \in \mathrm{ri}(\partial R(\bar{x}))$ is non-degenerate.*
- *$\mathcal{M}_{\bar{x}}$ is an affine subspace.*

In turn, $\mathrm{Id} + H_{\overline{R}}$ is invertible and $(\mathrm{Id} + H_{\overline{R}})^{-1}$ is symmetric positive definite with all eigenvalues in $]0, 1]$.

Together with the previous results, Lemma B.5 allows us to linearize the generalized proximal mapping defined below.

Definition B.1 (Generalised proximal mapping). Let $R \in \Gamma_0(\mathbb{R}^n)$ and $\gamma > 0$, the generalized proximal mapping of R is defined by

$$\mathrm{prox}^A_{\gamma R}(\bar{w}) \overset{\text{def}}{=} \mathrm{argmin}_{x \in \mathbb{R}^n} \gamma R(x) + \tfrac{1}{2}\|Ax - \bar{w}\|^2, \tag{B.5}$$

where $\bar{w} \in \mathbb{R}^p$ and $A \in \mathbb{R}^{p \times n}$ has full column rank.

Since A has full column rank, $\mathrm{prox}_{\gamma R}^{A}$ is a single-valued mapping. When $A = \mathrm{Id}$, (B.5) reduces to the standard definition of proximity operator. Denote $\bar{x} \overset{\text{def}}{=} \mathrm{prox}_{\gamma R}^{A}(\bar{w})$, owing to the optimality condition, we have $\bar{u} \overset{\text{def}}{=} -A^T(A\bar{x}-\bar{w})/\gamma \in \partial R(\bar{x})$. Suppose R is partly smooth at $\bar{x}$ relative to $\mathcal{M}_{\bar{x}}$, and let $\bar{R} \overset{\text{def}}{=} \gamma R(x) - \langle x,\, \gamma\bar{u}\rangle$ be a smooth perturbation of γR. Define $A_{T_{\bar{x}}} = A \circ \mathcal{P}_{T_{\bar{x}}}$ which also has full column rank, hence $A_{T_{\bar{x}}}^T A_{T_{\bar{x}}}$ is invertible and we define

$$M_{\bar{R}} \overset{\text{def}}{=} A_{T_{\bar{x}}}(\mathrm{Id} + (A_{T_{\bar{x}}}^T A_{T_{\bar{x}}})^{-1} H_{\bar{R}})^{-1} (A_{T_{\bar{x}}}^T A_{T_{\bar{x}}})^{-1} A_{T_{\bar{x}}}^T.$$

Lemma B.6 ([66, Lemma C.9]). *For the proximal mapping defined in* (B.5)*, suppose $R \in \Gamma_0(\mathbb{R}^n)$ is partly smooth at $\bar{x}$ relative to a C^2-smooth manifold $\mathcal{M}_{\bar{x}}$ and $\bar{u} \in \mathrm{ri}(\partial R(\bar{x}))$. Let $\{w_k\}_{k\in\mathbb{N}}$ be a sequence such that $w_k \to \bar{w}$ and $x_k = \mathrm{prox}_{\gamma R}^{A}(w_k) \to \bar{x}$, then for all k large enough, there hold $x_k \in \mathcal{M}_{\bar{x}}$ and*

$$A_{T_{\bar{x}}}(x_k - x_{k-1}) = M_{\bar{R}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|). \tag{B.6}$$

Remark B.7. When $A = \mathrm{Id}$, $\mathrm{prox}_{\gamma R}^{A}$ reduces to the standard proximal mapping and (B.6) simplifies to

$$x_k - x_{k-1} = \mathcal{P}_{T_{\bar{x}}}(\mathrm{Id} + H_{\bar{R}})^{-1}\mathcal{P}_{T_{\bar{x}}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|).$$

In [45] and references therein, to study the local linear convergence of first-order methods, linearization with respect to the limiting point $\bar{x}$ is provided, that is $x_k - \bar{x} = \mathcal{P}_{T_{\bar{x}}}(\mathrm{Id} + H_{\bar{R}})^{-1}\mathcal{P}_{T_{\bar{x}}}(w_k - \bar{w}) + o(\|w_k - \bar{w}\|)$.

For the sake of completeness, we provide the proof of Lemma B.6 below.

Proof. Since R is proper convex and lower semi-continuous, we have $R(x_k) \to R(\bar{x})$ and $\partial R(x_k) \ni u_k = -A^T(Ax_k - w_k)/\gamma \to \bar{u} \in \mathrm{ri}(\partial R(\bar{x}))$, hence $\mathrm{dist}(u_k, \partial R(\bar{x})) \to 0$. As a result, we have $x_k \in \mathcal{M}_{\bar{x}}$ owing to [39, Theorem 5.3] and $u_k \in \mathrm{ri}(\partial R(x_k))$ owing to [77] for all k large enough.

Denote $T_{x_k}, T_{x_{k-1}}$ the tangent spaces of $\mathcal{M}_{\bar{x}}$ at x_k and x_{k-1}. Denote $\tau_k : T_{x_k} \to T_{x_{k-1}}$ the parallel translation along the unique geodesic on $\mathcal{M}_{\bar{x}}$ joining x_k to x_{k-1}. From the definition of x_k, let $h_k = \gamma u_k$, we get

$$h_k \overset{\text{def}}{=} -A^T(Ax_k - w_k) \in \gamma\partial R(x_k) \quad \text{and} \quad h_{k-1} \overset{\text{def}}{=} -A^T(Ax_{k-1} - w_{k-1}) \in \gamma\partial R(x_{k-1}).$$

Projecting onto corresponding tangent spaces, applying Lemma B.1 and the parallel translation τ_k leads to

$$\begin{aligned}\gamma\tau_k\nabla_{\mathcal{M}_{\bar{x}}}R(x_k) &= \tau_k\mathcal{P}_{T_{x_k}}(h_k) = \mathcal{P}_{T_{x_{k-1}}}(h_k) + \big(\tau_k\mathcal{P}_{T_{x_k}} - \mathcal{P}_{T_{x_{k-1}}}\big)(h_k),\\ \gamma\nabla_{\mathcal{M}_{\bar{x}}}R(x_{k-1}) &= \mathcal{P}_{T_{x_{k-1}}}(h_{k-1}).\end{aligned}$$

The difference of the above two equalities yields

$$\gamma\tau_k\nabla_{\mathcal{M}_{\bar{x}}}R(x_k) - \gamma\nabla_{\mathcal{M}_{\bar{x}}}R(x_{k-1}) - \big(\tau_k\mathcal{P}_{T_{x_k}} - \mathcal{P}_{T_{x_{k-1}}}\big)(h_{k-1}) = \mathcal{P}_{T_{x_{k-1}}}(h_k - h_{k-1}) + \big(\tau_k\mathcal{P}_{T_{x_k}} - \mathcal{P}_{T_{x_{k-1}}}\big)(h_k - h_{k-1}). \tag{B.7}$$

Owing to the monotonicity of sub-differential, *i.e.* $\langle h_k - h_{k-1}, x_k - x_{k-1}\rangle \geq 0$, we get

$$\langle A^TA(x_k - x_{k-1}), x_k - x_{k-1}\rangle \leq \langle A^T(w_k - w_{k-1}), x_k - x_{k-1}\rangle \leq \|A\|\|w_k - w_{k-1}\|\|x_k - x_{k-1}\|.$$

Since A has full column rank, A^TA is symmetric positive definite, and there exists $\kappa > 0$ such that $\kappa\|x_k - x_{k-1}\|^2 \leq \langle A^TA(x_k - x_{k-1}), x_k - x_{k-1}\rangle$. Back to the above inequality, we get $\|x_k - x_{k-1}\| \leq \frac{\|A\|}{\kappa}\|w_k - w_{k-1}\|$. Therefore for $\|h_k - h_{k-1}\|$, we get

$$\|h_k - h_{k-1}\| = \|A^T(Ax_k - w_k) - A^T(Ax_{k-1} - w_{k-1})\| \leq \|A\|^2\|x_k - x_{k-1}\| + \|A\|\|w_k - w_{k-1}\| \leq \big(\tfrac{\|A\|^3}{\kappa} + \|A\|\big)\|w_k - w_{k-1}\|.$$

As a result, owing to Lemma B.3, we have for the term $(\tau_k\mathcal{P}_{T_{x_k}} - \mathcal{P}_{T_{x_{k-1}}})(h_k - h_{k-1})$ in (B.7) that

$$\big(\tau_k\mathcal{P}_{T_{x_k}} - \mathcal{P}_{T_{x_{k-1}}}\big)(h_k - h_{k-1}) = o(\|h_k - h_{k-1}\|) = o(\|w_k - w_{k-1}\|).$$

Define $\bar{R}_{k-1}(x) \overset{\text{def}}{=} \gamma R(x) - \langle x,\, h_{k-1}\rangle$ and $H_{\bar{R},k-1} \overset{\text{def}}{=} \mathcal{P}_{T_{x_{k-1}}} \nabla^2_{\mathcal{M}_{\bar{x}}} \bar{R}_{k-1}(x_{k-1}) \mathcal{P}_{T_{x_{k-1}}}$, then with Lemma B.4 the Riemannian Taylor expansion, we have for the first line of (B.7)

$$\begin{aligned}
\gamma\tau_k \nabla_{\mathcal{M}_{\bar{x}}} R(x_k) - \gamma\nabla_{\mathcal{M}_{\bar{x}}} R(x_{k-1}) - \big(\tau_k \mathcal{P}_{T_{x_k}} - \mathcal{P}_{T_{x_{k-1}}}\big)(h_{k-1}) &= \tau_k\big(\gamma\nabla_{\mathcal{M}_{\bar{x}}} R(x_k) - \mathcal{P}_{T_{x_k}}(h_{k-1})\big) - \big(\gamma\nabla_{\mathcal{M}_{\bar{x}}} R(x_{k-1}) - \mathcal{P}_{T_{x_{k-1}}}(h_{k-1})\big) \\
&= \tau_k \nabla_{\mathcal{M}_{\bar{x}}} \bar{R}_{k-1}(x_k) - \nabla_{\mathcal{M}_{\bar{x}}} \bar{R}_{k-1}(x_{k-1}) \\
&= H_{\bar{R},k-1}(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|) \\
&= H_{\bar{R},k-1}(x_k - x_{k-1}) + o(\|w_k - w_{k-1}\|).
\end{aligned} \tag{B.8}$$

Back to (B.7), we get

$$H_{\bar{R},k-1}(x_k - x_{k-1}) = \mathcal{P}_{T_{x_{k-1}}}(h_k - h_{k-1}) + o(\|w_k - w_{k-1}\|). \tag{B.9}$$

Define $\bar{R}(x) \overset{\text{def}}{=} \gamma R(x) - \langle x,\, \bar{h}\rangle$ and $H_{\bar{R}} = \mathcal{P}_{T_{\bar{x}}} \nabla^2_{\mathcal{M}_{\bar{x}}} \bar{R}(\bar{x}) \mathcal{P}_{T_{\bar{x}}}$, then from (B.9) that

$$H_{\bar{R}}(x_k - x_{k-1}) + \big(H_{\bar{R},k-1} - H_{\bar{R}}\big)(x_k - x_{k-1}) = \mathcal{P}_{T_{\bar{x}}}(h_k - h_{k-1}) + \big(\mathcal{P}_{T_{x_{k-1}}} - \mathcal{P}_{T_{\bar{x}}}\big)(h_k - h_{k-1}) + o(\|w_k - w_{k-1}\|). \tag{B.10}$$

Owing to continuity, we have $H_{\bar{R},k-1} \to H_{\bar{R}}$ and $\mathcal{P}_{T_{x_{k-1}}} \to \mathcal{P}_{T_{\bar{x}}}$,

$$\begin{aligned}
\lim_{k\to+\infty} \tfrac{\|(H_{\bar{R},k-1} - H_{\bar{R}})(x_k - x_{k-1})\|}{\|x_k - x_{k-1}\|} &\leq \lim_{k\to+\infty} \tfrac{\|H_{\bar{R},k-1} - H_{\bar{R}}\|\|x_k - x_{k-1}\|}{\|x_k - x_{k-1}\|} = \lim_{k\to+\infty} \|H_{\bar{R},k-1} - H_{\bar{R}}\| = 0, \\
\lim_{k\to+\infty} \tfrac{\|(\mathcal{P}_{T_{x_{k-1}}} - \mathcal{P}_{T_{\bar{x}}})(w_k - w_{k-1})\|}{\|w_k - w_{k-1}\|} &\leq \lim_{k\to+\infty} \tfrac{\|\mathcal{P}_{T_{x_{k-1}}} - \mathcal{P}_{T_{\bar{x}}}\|\|w_k - w_{k-1}\|}{\|w_k - w_{k-1}\|} = \lim_{k\to+\infty} \|\mathcal{P}_{T_{x_{k-1}}} - \mathcal{P}_{T_{\bar{x}}}\| = 0,
\end{aligned}$$

and $\lim_{k\to+\infty} \frac{\|(\mathcal{P}_{T_{x_{k-1}}} - \mathcal{P}_{T_{\bar{x}}})(x_k - x_{k-1})\|}{\|x_k - x_{k-1}\|} = 0$. Combining this with the definition of u_k, the fact that $x_k - x_{k-1} = \mathcal{P}_{T_{\bar{x}}}(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|)$ from Lemma B.2, and denoting $A_{T_{\bar{x}}} = A \circ \mathcal{P}_{T_{\bar{x}}}$, equation (B.10) can be written as

$$\begin{aligned}
H_{\bar{R}}(x_k - x_{k-1}) = \mathcal{P}_{T_{\bar{x}}}(u_k - u_{k-1}) + o(\|w_k - w_{k-1}\|) &= -\mathcal{P}_{T_{\bar{x}}}(A^T(Ax_k - w_k) - A^T(Ax_{k-1} - w_{k-1})) + o(\|w_k - w_{k-1}\|) \\
&= -\mathcal{P}_{T_{\bar{x}}} A^T A(x_k - x_{k-1}) + \mathcal{P}_{T_{\bar{x}}} A^T(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|) \\
&= -A^T_{T_{\bar{x}}} A_{T_{\bar{x}}}(x_k - x_{k-1}) + A^T_{T_{\bar{x}}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|).
\end{aligned} \tag{B.11}$$

Since A has full rank, so is $A_{T_{\bar{x}}}$. Hence $A^T_{T_{\bar{x}}} A_{T_{\bar{x}}}$ is invertible and from above we have

$$\big(\mathrm{Id} + (A^T_{T_{\bar{x}}} A_{T_{\bar{x}}})^{-1} H_{\bar{R}}\big)(x_k - x_{k-1}) = (A^T_{T_{\bar{x}}} A_{T_{\bar{x}}})^{-1} A^T_{T_{\bar{x}}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|).$$

Denote $M_{\bar{R}} = A_{T_{\bar{x}}}(\mathrm{Id} + (A^T_{T_{\bar{x}}} A_{T_{\bar{x}}})^{-1} H_{\bar{R}})^{-1} (A^T_{T_{\bar{x}}} A_{T_{\bar{x}}})^{-1} A^T_{T_{\bar{x}}}$, then

$$A_{T_{\bar{x}}}(x_k - x_{k-1}) = M_{\bar{R}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|), \tag{B.12}$$

which concludes the proof. □

B.3 Proof of main propositions

B.3.1 Forward–Backward splitting method

Proof of Proposition 3.5. The proof of the result is split into several steps.

Linearization of Forward–Backward splitting The convergence of $\{x_k\}_{k\in\mathbb{N}}$ to a global minimizer $x^\star \in \mathrm{Argmin}(\Phi)$ can be guaranteed under proper choices of a_k and γ, we refer to [47] and the reference therein for detailed discussion. When $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star})$ and the non-degeneracy condition ($\mathrm{ND_{FB}}$) holds, then there exists $K > 0$ such that for all $k \geq K$, there holds $x_k \in \mathcal{M}_{x^\star}$; see [47, Theorem 3.4].

Denote $u^\star = -\nabla F(x^\star)$, from the non-degeneracy condition ($\mathrm{ND_{FB}}$) we have $u^\star \in \mathrm{ri}(\partial R(x^\star))$. Define $\bar{R}(x) = \gamma R(x) - \langle x,\, -\nabla F(x^\star)\rangle$, and the following matrices

$$H_{\bar{R}} \overset{\text{def}}{=} \mathcal{P}_{T_{x^\star}} \nabla^2_{\mathcal{M}_{x^\star}} \bar{R}(x^\star) \mathcal{P}_{T_{x^\star}} \quad \text{and} \quad M_{\bar{R}} \overset{\text{def}}{=} \mathcal{P}_{T_{x^\star}}(\mathrm{Id} + H_{\bar{R}})^{-1} \mathcal{P}_{T_{x^\star}}.$$

Let $w_k \overset{\text{def}}{=} x_k - \gamma\nabla F(x_k)$, then the update of x_k entails that $x_{k+1} = \mathrm{argmin}_{x\in\mathbb{R}^n} \gamma R(x) + \frac{1}{2}\|x - w_k\|^2$. Owing to Lemma B.6, we get that

$$x_{k+1} - x_k = M_{\bar{R}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|). \tag{B.13}$$

Next we deal with the term $w_k - w_{k-1}$, from the local C^2-smoothness of F we get

$$\begin{aligned} w_k - w_{k-1} &= x_k - x_{k-1} - \gamma\big(\nabla F(x_k) - \gamma\nabla F(x_{k-1})\big) \\ &= x_k - x_{k-1} - \gamma\nabla^2 F(x_{k-1})(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|) \\ &= x_k - x_{k-1} - \gamma\nabla^2 F(x^\star)(x_k - x_{k-1}) - \gamma\big(\nabla^2 F(x_{k-1}) - \nabla^2 F(x^\star)\big)(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|). \end{aligned}$$

Since $x_k \to x^\star$, we have $\nabla^2 F(x_{k-1}) \to \nabla^2 F(x^\star)$, then from above we get

$$w_k - w_{k-1} = \big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big)(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|).$$

As $\mathrm{Id} - \gamma\nabla F$ is non-expansive, then $\|w_k - w_{k-1}\| \leq \|x_k - x_{k-1}\|$. Therefore, back to (B.13),

$$x_{k+1} - x_k = M_{\bar{R}}\big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big)(x_k - x_{k-1}) + o(\|x_k - x_{k-1}\|), \tag{B.14}$$

which is the desired linearization with $M_{\mathrm{FB}} = M_{\bar{R}}\big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big)$.

Spectral properties of M_{FB} In this part, we briefly discuss the spectrum of M_{FB} where more detailed accountant can be found in [47] and [45, Chapter 6]. Owing to Lemma B.5, we have $H_{\bar{R}}$ is symmetric positive semi-definite, hence $M_{\bar{R}}$ is symmetric positive definite with all its eigenvalues in $]0,1]$. As a result, we have

$$M_{\bar{R}}\big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big) = M_{\bar{R}}^{1/2} M_{\bar{R}}^{1/2}\big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big) M_{\bar{R}}^{1/2} M_{\bar{R}}^{-1/2} \sim M_{\bar{R}}^{1/2}\big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big) M_{\bar{R}}^{1/2}, \tag{B.15}$$

where $M_{\bar{R}}^{1/2}\big(\mathrm{Id} - \gamma\nabla^2 F(x^\star)\big) M_{\bar{R}}^{1/2}$ is symmetric, with all its eigenvalues in $]-1,1]$ since $\gamma \in]0, 2/L[$.

Trajectory of Forward–Backward splitting Let $1 \geq e_1 > e_2$ be the two largest eigenvalues of Γ, and suppose $a \leq \frac{(1-\sqrt{1-e_2})^2}{e_2}$, then the leading eigenvalue of Σ is real which is denoted by σ. Then

$$\cos(\theta_k) = \frac{\langle v_k, v_{k-1}\rangle}{\|v_k\|\|v_{k-1}\|} = \frac{\langle Mv_{k-1}, v_{k-1}\rangle}{\|v_{k-1}\|^2}\frac{\|v_{k-1}\|}{\|Mv_{k-1}\|} = \frac{\langle Mv_{k-1}, v_{k-1}\rangle}{\|v_{k-1}\|^2}\frac{\|v_{k-1}\|}{\|Mv_{k-1}\|}.$$

Note that $\frac{\langle Mv_{k-1}, v_{k-1}\rangle}{\|v_{k-1}\|^2}$ is simply the power iteration of M with starting vector v_0 and converges to σ_1 with $\frac{\langle Mv_{k-1}, v_{k-1}\rangle}{\|v_{k-1}\|^2} = \sigma + o(1)$, $\frac{\|v_{k-1}\|}{\|Mv_{k-1}\|} \to 1/\sigma$. Combine them together we get $\cos(\theta_k) \to 1$, hence $\theta_k \to 0$. □

B.3.2 Douglas–Rachford splitting and ADMM

Proof of Theorems 3.8 & 3.9. The proof is also divided into parts.

Linearization of Douglas–Rachford Owing to [48, Theorem 5.1], we have that when $R \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star}^R)$, $J \in \mathrm{PSF}_{x^\star}(\mathcal{M}_{x^\star}^J)$ and the non-degeneracy condition ($\mathrm{ND}_{\mathrm{DR}}$) holds, there exists $K > 0$ such that for all $k \geq K$, $(u_k, x_k) \in \mathcal{M}_{x^\star}^R \times \mathcal{M}_{x^\star}^J$.

From (3.3), the update of x_k: define $\bar{J}(x) \overset{\mathrm{def}}{=} \gamma J(x) - \langle x, z^\star - x^\star\rangle$, $H_{\bar{J}} \overset{\mathrm{def}}{=} \mathcal{P}_{T_{x^\star}^J}\nabla^2_{\mathcal{M}_{x^\star}^J}\bar{J}(x^\star)\mathcal{P}_{T_{x^\star}^J}$ and $M_{\bar{J}} = \mathcal{P}_{T_{x^\star}^J}(\mathrm{Id} + H_{\bar{J}})^{-1}\mathcal{P}_{T_{x^\star}^J}$, then from Lemma B.6 we get

$$x_k - x_{k-1} = M_{\bar{J}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|). \tag{B.16}$$

Now for u_{k+1}, let $w_k = 2x_k - z_k$, we get from Lemma B.6 that

$$u_{k+1} - u_k = M_{\bar{R}}(w_k - w_{k-1}) + o(\|w_k - w_{k-1}\|),$$

Define $\bar{R}(x) \overset{\mathrm{def}}{=} \gamma R(x) - \langle x, x^\star - z^\star\rangle$, $H_{\bar{R}} = \mathcal{P}_{T_{x^\star}^R}\nabla^2_{\mathcal{M}_{x^\star}^R}\bar{R}(x^\star)\mathcal{P}_{T_{x^\star}^R}$ and $M_{\bar{R}} = \mathcal{P}_{T_{x^\star}^R}(\mathrm{Id} + H_{\bar{R}})^{-1}\mathcal{P}_{T_{x^\star}^R}$. Since $\|x_k - x_{k-1}\| \leq \|z_k - z_{k-1}\|$, we get from above that

$$u_{k+1} - u_k = 2M_{\bar{R}}M_{\bar{J}}(z_k - z_{k-1}) - M_{\bar{R}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|). \tag{B.17}$$

Summing up (B.16) and (B.17), we get

$$\begin{aligned} z_{k+1} - z_k &= (z_k + u_{k+1} - x_k) - (z_{k-1} + u_k - x_{k-1}) \\ &= (z_k - z_{k-1}) + (u_{k+1} - u_k) - (x_k - x_{k-1}) \\ &= (\mathrm{Id} + 2M_{\bar{R}}M_{\bar{J}} - M_{\bar{R}} - M_{\bar{J}})(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|) \\ &= M_{\mathrm{DR}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|). \end{aligned} \tag{B.18}$$

Owing to Lemma B.5, we have $H_{\bar{R}}, H_{\bar{J}}$ are symmetric positive semi-definite, hence maximal monotone, consequently $(\mathrm{Id}+H_{\bar{R}})^{-1}, (\mathrm{Id}+H_{\bar{J}})^{-1}$ are firmly non-expansive. Since $\mathcal{P}_{T^R_{x^\star}}, \mathcal{P}_{T^J_{x^\star}}$ are projection operators, thens firmly non-expansive. As a result, both $M_{\bar{R}} = \mathcal{P}_{T^R_{x^\star}}(\mathrm{Id}+H_{\bar{R}})^{-1}\mathcal{P}_{T^R_{x^\star}}, M_{\bar{J}} = \mathcal{P}_{T^J_{x^\star}}(\mathrm{Id}+H_{\bar{J}})^{-1}\mathcal{P}_{T^J_{x^\star}}$ are firmly non-expansive owing to [7, Example 4.14], and M_{DR} is firmly non-expansive [7, Proposition 4.31].

Spectral properties of M_{DR} Here we present a brief summary on the spectral properties of M_{DR} and refer to [48, 10] and the reference therein for detailed analysis of the spectral properties of M_{DR}. When R, J are locally polyhedral around $x^\star$, then $H_{\bar{R}}, H_{\bar{J}}$ vanish and consequently

$$M_{\bar{R}} = \mathcal{P}_{T^R_{x^\star}}, \;\; M_{\bar{J}} = \mathcal{P}_{T^J_{x^\star}} \quad \text{and} \quad M_{\mathrm{DR}} = \mathcal{P}_{T^R_{x^\star}}\mathcal{P}_{T^J_{x^\star}} + \mathcal{P}_{S^R_{x^\star}}\mathcal{P}_{S^J_{x^\star}}$$

where $S^R_{x^\star} = (T^R_{x^\star})^\perp$ and $S^J_{x^\star} = (T^J_{x^\star})^\perp$. Denote the dimension of $T^R_{x^\star}, T^J_{x^\star}$ are $\dim(T^R_{x^\star}) = p, \dim(T^J_{x^\star}) = q$. Without the loss of generality, we assume that $1 \leq p \leq q \leq n, v$ and $\dim(T^R_{x^\star} \cap T^J_{x^\star}) = d$. Consequently, there are $r = p - d$ principal angles $(\zeta_i)_{i=1,\ldots,r}$ between $T^R_{x^\star}$ and $T^J_{x^\star}$ that are strictly greater than 0 and smaller than $\pi/2$. Suppose that $\zeta_1 \leq \cdots \leq \zeta_r$. Define the following two diagonal matrices

$$C = \mathrm{diag}\big(\cos(\zeta_1), \cdots, \cos(\zeta_r)\big) \quad \text{and} \quad S = \mathrm{diag}\big(\sin(\zeta_1), \cdots, \sin(\zeta_r)\big).$$

Owing to [10, 29], there exists a real orthogonal matrix U such that

$$M_{\mathrm{DR}} = U \left[\begin{array}{cc|cc} C^2 & CS & 0 & 0 \\ -CS & C^2 & 0 & 0 \\ \hline 0 & 0 & 0_{q-p+2d} & 0 \\ 0 & 0 & 0 & \mathrm{Id}_{n-p-q} \end{array}\right] U^T,$$

which indicates M_{DR} is normal and all its eigenvalues are inside unit disc.

Trajectory of Douglas–Rachford The above spectral properties of M_{DR} indicates that M_{DR} is a Type II matrix. Let $M^\infty_{\mathrm{DR}} = \lim_{k\to+\infty} M^k_{\mathrm{DR}}$ and $\widetilde{M}_{\mathrm{DR}} = M_{\mathrm{DR}} - M^\infty_{\mathrm{DR}}$, then we have

$$\widetilde{M}_{\mathrm{DR}} = U \left[\begin{array}{cc|c} C^2 & CS & 0 \\ -CS & C^2 & 0 \\ \hline 0 & 0 & 0_{n-2r} \end{array}\right] U^T.$$

Denote $\theta_F = \zeta_1$ the Friedrichs angle between $T^R_{x^\star}$ and $T^J_{x^\star}$, then invoking Proposition A.4 we obtain the trajectory of Douglas–Rachford splitting method. □

Proof of Proposition 3.12. From the above proof, we have for x_k that $x_k - x_{k-1} = M_{\bar{J}}(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|)$ Lemma B.2. Now for u_{k+1}, since R is smooth differentiable, we have

$$2x_k - z_k - u_{k+1} = \gamma\nabla R(u_{k+1}) \quad \text{and} \quad 2x_{k-1} - z_{k-1} - u_k = \gamma\nabla R(u_k).$$

Since we assume that R is locally C^2-smooth around $x^\star$, then when u_{k+1}, u_k is close enough, *i.e.* for sufficiently large k,

$$\begin{aligned}
(2x_k - z_k - u_{k+1}) - (2x_{k-1} - z_{k-1} - u_k) &= \gamma\nabla R(u_{k+1}) - \gamma\nabla R(u_k) \\
&= \gamma\nabla^2 R(u_k)(u_{k+1} - u_k) + o(\|u_{k+1} - u_k\|) \\
&= \gamma\nabla^2 R(u^\star)(u_{k+1} - u_k) + \gamma\big(\nabla^2 R(u_k) - \nabla^2 R(u^\star)\big)(u_{k+1} - u_k) + o(\|u_{k+1} - u_k\|) \\
&= \gamma\nabla^2 R(u^\star)(u_{k+1} - u_k) + o(\|u_{k+1} - u_k\|) \\
&= \gamma\nabla^2 R(u^\star)(u_{k+1} - u_k) + o(\|z_k - z_{k-1}\|),
\end{aligned}$$

and consequently, let $M_R = (\mathrm{Id} + \gamma\nabla^2 R(u^\star))^{-1}$,

$$\begin{aligned}
u_{k+1} - u_k &= M_R\big(2(x_k - x_{k-1}) - (z_k - z_{k-1})\big) + o(\|z_k - z_{k-1}\|) \\
&= 2M_R(x_k - x_{k-1}) - M_R(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|) \\
&= 2M_R M_{\bar{J}}(x_k - x_{k-1}) - M_R(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|).
\end{aligned}$$

Summing up we get

$$\begin{aligned} z_{k+1} - z_k &= (z_k - z_{k-1}) + (u_{k+1} - u_k) - (x_k - x_{k-1}) \\ &= (\mathrm{Id} + 2M_R M_{\bar{J}} - M_R - M_{\bar{J}})(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|) \\ &= \big(\tfrac{1}{2}\mathrm{Id} + \tfrac{1}{2}(2M_R - \mathrm{Id})(2M_{\bar{J}} - \mathrm{Id})\big)(z_k - z_{k-1}) + o(\|z_k - z_{k-1}\|). \end{aligned}$$

Now we discuss the spectral property of M_{DR}. Owing to convexity, we have $\nabla^2 R(u^\star)$ is symmetric positive semi-definite, hence maximal monotone and M_R is firmly non-expansive. When $\gamma < \frac{1}{\|\nabla^2 R(x^\star)\|}$, we have that all the eigenvalues of M_R are in $]1/2, 1]$, consequently $W_R \overset{\mathrm{def}}{=} 2M_R - \mathrm{Id}$ is symmetric positive definite. Therefore, we get

$$\begin{aligned} \tfrac{1}{2}\mathrm{Id} + \tfrac{1}{2}W_R\big(2M_{\bar{J}} - \mathrm{Id}\big) &= W_R^{1/2}\big(\tfrac{1}{2}\mathrm{Id} + \tfrac{1}{2}W_R^{1/2}\big(2M_{\bar{J}} - \mathrm{Id}\big)W_R^{1/2}\big)W_R^{-1/2} \\ &\sim \tfrac{1}{2}\mathrm{Id} + \tfrac{1}{2}W_R^{1/2}\big(2M_{\bar{J}} - \mathrm{Id}\big)W_R^{1/2}, \end{aligned}$$

and $\frac{1}{2}\mathrm{Id} + \frac{1}{2}W_R^{1/2}(2M_{\bar{J}} - \mathrm{Id})W_R^{1/2}$ is symmetric positive semi-definite with all it eigenvalues in $[0, 1]$. □

B.3.3 Primal–Dual splitting

Proof of Theorems 3.14 & 3.15. Similar to above, the proof is divided into parts.

Linearization of Primal–Dual From the x_k in (3.5), define $\bar{R}(x) \overset{\mathrm{def}}{=} \gamma_R R(x) - \langle x, -\gamma_R L^T w^\star\rangle$ and $H_{\bar{R}} \overset{\mathrm{def}}{=} \mathcal{P}_{T^R_{x^\star}} \nabla^2_{\mathcal{M}^R_{x^\star}} \bar{R}(x^\star)\mathcal{P}_{T^R_{x^\star}}$. Applying Lemma B.6 we get

$$x_{k+1} - x_k = M_{\bar{R}}(x_k - x_{k-1}) - \gamma_R M_{\bar{R}}\bar{L}^T(w_k - w_{k-1}) + o(\|x_k - x_{k-1}\| + \gamma_R\|L\|\|w_k - w_{k-1}\|). \tag{B.19}$$

Now for the update of w_{k+1}. Since $\tau \in [0, 1]$, we have

$$\begin{aligned} \|\bar{x}_{k+1} - \bar{x}_k\| \leq (1+\tau)\|x_{k+1} - x_k\| + \tau\|x_k - x_{k-1}\| &\leq 4(\|x_k - x_{k-1}\| + \gamma_R\|L\|\|w_k - w_{k-1}\|) + \|x_k - x_{k-1}\| \\ &= 5\|x_k - x_{k-1}\| + 4\gamma_R\|L\|\|w_k - w_{k-1}\|. \end{aligned}$$

Define $\overline{J^*}(w) \overset{\mathrm{def}}{=} \gamma_J J^*(w) - \langle w, \gamma_J L x^\star\rangle$ and $H_{\overline{J^*}} = \mathcal{P}_{T^{J^*}_{w^\star}} \nabla^2_{\mathcal{M}^{J^*}_{w^\star}} \overline{J^*}(w^\star)\mathcal{P}_{T^{J^*}_{w^\star}}$, applying Lemma B.6 then yields

$$\begin{aligned} w_{k+1} - w_k &= M_{\overline{J^*}}(w_k - w_{k-1}) + \gamma_J M_{\overline{J^*}}\bar{L}(\bar{x}_{k+1} - \bar{x}_k) + \text{small } o\text{-terms} \\ &= M_{\overline{J^*}}(w_k - w^\star) + (1+\tau)\gamma_J M_{\overline{J^*}}\bar{L}(x_{k+1} - x_k) - \tau\gamma_J M_{\overline{J^*}}\bar{L}(x_k - x_{k-1}) + \text{small } o\text{-terms} \\ &= M_{\overline{J^*}}(w_k - w_{k-1}) - \tau\gamma_J M_{\overline{J^*}}\bar{L}(x_k - x_{k-1}) \\ &\quad + (1+\tau)\gamma_J M_{\overline{J^*}}\bar{L}\big(M_{\bar{R}}(x_k - x_{k-1}) - \gamma_R M_{\bar{R}}\bar{L}^T(w_k - w_{k-1})\big) + \text{small } o\text{-terms} \\ &= \big(M_{\overline{J^*}} - (1+\tau)\gamma_J\gamma_R M_{\overline{J^*}}\bar{L}M_{\bar{R}}\bar{L}^T\big)(w_k - w_{k-1}) + \big((1+\tau)\gamma_J M_{\overline{J^*}}\bar{L}M_{\bar{R}} - \tau\gamma_J M_{\overline{J^*}}\bar{L}\big)(x_k - x_{k-1}) \\ &\quad + o(\|w_{k+1} - w_k\|) + o(\|x_{k+1} - x_k\|) \\ &\quad + o(\|x_k - x_{k-1}\| + \gamma_R\|L\|\|w_k - w_{k-1}\|) + o(\|w_k - w_{k-1}\| + \gamma_J\|L\|\|x_k - x_{k-1}\|). \end{aligned} \tag{B.20}$$

Combining (B.19) and (B.20), we get

$$\begin{aligned} \begin{pmatrix} x_{k+1} - x_k \\ w_{k+1} - w_k \end{pmatrix} = &\begin{bmatrix} M_{\bar{R}} & -\gamma_R M_{\bar{R}}\bar{L}^T \\ (1+\tau)\gamma_J M_{\overline{J^*}}\bar{L}M_{\bar{R}} - \tau\gamma_J M_{\overline{J^*}}\bar{L} & M_{\overline{J^*}} - (1+\tau)\gamma_J\gamma_R M_{\overline{J^*}}\bar{L}M_{\bar{R}}\bar{L}^T \end{bmatrix} \begin{pmatrix} x_k - x_{k-1} \\ w_k - w_{k-1} \end{pmatrix} \\ &+ o(\|w_{k+1} - w_k\|) + o(\|x_{k+1} - x_k\|) + o(\|x_k - x_{k-1}\| + \gamma_R\|L\|\|w_k - w_{k-1}\|) \\ &+ o(\|w_k - w_{k-1}\| + \gamma_J\|L\|\|x_k - x_{k-1}\|) \end{aligned} \tag{B.21}$$

Now we consider the small o-terms. Let a_1, a_2 be two constants, then we have

$$|a_1| + |a_2| = \sqrt{(|a_1| + |a_2|)^2} \leq \sqrt{2(a_1^2 + a_2^2)} = \sqrt{2}\left\|\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\right\|.$$

Denote $b = \max\{1, \gamma_J\|L\|, \gamma_R\|L\|\}$, then

$$\begin{aligned} &(\|w_k - w_{k-1}\| + \gamma_J\|L\|\|x_k - x_{k-1}\|) + (\|x_k - x_{k-1}\| + \gamma_R\|L\|\|w_k - w_{k-1}\|) \\ &\leq 2b(\|x_k - x_{k-1}\| + \|w_k - w_{k-1}\|) \leq 2\sqrt{2}b\left\|\begin{pmatrix} x_k - x_{k-1} \\ w_k - w_{k-1} \end{pmatrix}\right\|. \end{aligned}$$

Then for $\|w_{k+1}-w_k\|+\|x_{k+1}-x_k\|$, we have

$$\|w_{k+1}-w_k\|+\|x_{k+1}-x_k\| \leq \sqrt{2}\left\|\begin{pmatrix} x_{k+1}-x_k \\ w_{k+1}-w_k \end{pmatrix}\right\|.$$

Combining these into the small o-terms of (B.21), we obtain

$$z_{k+1}-z_k = M_{\mathrm{PD}}(z_k-z_{k-1})+o(\|z_k-z_{k-1}\|).$$

Spectral properties of M_{PD} We refer to [49, Proposition 3.5] about the non-expansiveness of M_{PD}, below we provide the spectral analysis of M_{PD} for the case when both R, J^* are locally polyhedral around $x^\star$ and $w^\star$ respectively, the analysis can also be found in [49].

When R, J are locally polyhedral around $x^\star$, then $H_{\bar{R}}, H_{\bar{J}}$ vanish and consequently

$$M_{\bar{R}} = \mathrm{Id}_n,\ \ M_{\bar{J}} = \mathrm{Id}_m \quad \text{and} \quad M_{\mathrm{PD}} = \begin{bmatrix} \mathrm{Id}_n & -\gamma_{_R}\bar{L}^T \\ \gamma_{_J}\bar{L} & \mathrm{Id}_m-(1+\tau)\gamma_{_J}\gamma_{_R}\bar{L}\bar{L}^T \end{bmatrix}.$$

Let $p \overset{\text{def}}{=} \dim(T^R_{x^\star}), q \overset{\text{def}}{=} \dim(T^{J^*}_{w^\star})$ be the dimensions of $T^R_{x^\star}$ and $T^{J^*}_{w^\star}$ respectively, define $S^R_{x^\star} = (T^R_{x^\star})^\perp$ and $S^{J^*}_{w^\star} = (T^{J^*}_{w^\star})^\perp$. Assume that $q \geq p$, where as the other direction can be treated similarly. Let $\bar{L} = X\Sigma_{\bar{L}}Y^T$ the singular value decomposition of $\bar{L}$, denote the rank of $\bar{L}$ as $l \overset{\text{def}}{=} \mathrm{rank}(\bar{L})$. Clearly, we have $l \leq p$. With the SVD of $\bar{L}$, for M_{PD}, we have

$$\begin{aligned} M_{\mathrm{PD}} = \begin{bmatrix} \mathrm{Id}_n & -\gamma_{_R}\bar{L}^T \\ \gamma_{_J}\bar{L} & \mathrm{Id}_m-(1+\tau)\gamma_{_R}\gamma_{_J}\bar{L}\bar{L}^T \end{bmatrix} &= \begin{bmatrix} YY^T & -\gamma_{_R}Y\Sigma^*_{\bar{L}}X^T \\ \gamma_{_J}X\Sigma_{\bar{L}}Y^T & XX^T-(1+\tau)\gamma_{_R}\gamma_{_J}X\Sigma^2_{\bar{L}}X^T \end{bmatrix} \\ &= \begin{bmatrix} Y & \\ & X \end{bmatrix} \underbrace{\begin{bmatrix} \mathrm{Id}_n & -\gamma_{_R}\Sigma^*_{\bar{L}} \\ \gamma_{_J}\Sigma_{\bar{L}} & \mathrm{Id}_m-(1+\tau)\gamma_{_R}\gamma_{_J}\Sigma^2_{\bar{L}} \end{bmatrix}}_{W} \begin{bmatrix} Y^T & \\ & X^T \end{bmatrix}. \end{aligned} \tag{B.22}$$

Since we assume that $\mathrm{rank}(\bar{L}) = l \leq p$, then $\Sigma_{\bar{L}}$ can be represented as $\Sigma_{\bar{L}} = \begin{bmatrix} \Sigma_l & 0_{l,n-l} \\ 0_{m-l,l} & 0_{m-l,n-l} \end{bmatrix}$ where $\Sigma_l = (\sigma_j)_{j=1,...,l}$. Back to W, we have there exists an elementary transformation E such that

$$W = \begin{bmatrix} \mathrm{Id}_l & 0_{l,n-l} & -\gamma_{_R}\Sigma_l & 0_{l,m-l} \\ 0_{n-l,l} & \mathrm{Id}_{n-l} & 0_{n-l,l} & 0_{n-l,m-l} \\ \gamma_{_J}\Sigma_l & 0_{l,n-l} & \mathrm{Id}_l-(1+\tau)\gamma_{_R}\gamma_{_J}\Sigma^2_l & 0_{l,m-l} \\ 0_{m-l,l} & 0_{m-l,n-l} & 0_{m-l,l} & \mathrm{Id}_{m-l} \end{bmatrix} = E \begin{bmatrix} \mathrm{Id}_l & -\gamma_{_R}\Sigma_l & 0_{l,m+n-2l} \\ \gamma_{_J}\Sigma_l & \mathrm{Id}_l-(1+\tau)\gamma_{_R}\gamma_{_J}\Sigma^2_l & 0_{l,m+n-2l} \\ 0_{m+n-2l,l} & 0_{m+n-2l,l} & \mathrm{Id}_{m+n-2l} \end{bmatrix} E.$$

Clearly, 1 is an eigenvalue of M_{PD} with multiplicity $m+n-2l$. Next we deal with the block diagonal matrix

$$D = \begin{bmatrix} \mathrm{Id}_l & -\gamma_{_R}\Sigma_l \\ \gamma_{_J}\Sigma_l & \mathrm{Id}_l-(1+\tau)\gamma_{_R}\gamma_{_J}\Sigma^2_l \end{bmatrix}.$$

Again, there exists another elementary transformation E' such that

$$D = E' \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_l, \end{bmatrix} E',$$

where for each $D_i, i = 1, ..., l$, we have $D_i = \begin{bmatrix} 1 & -\gamma_{_R}\sigma_i \\ \gamma_{_J}\sigma_i & 1-(1+\tau)\gamma_{_R}\gamma_{_J}\sigma^2_i \end{bmatrix}$, which is the 2×2 matrix studied in Type III matrix. Therefore, for each $i = 1, ..., l$, the eigenvalue of D_i is

$$\rho_i = \frac{(2-(1+\tau)\gamma_{_J}\gamma_{_R}\sigma^2_j) \pm \sqrt{(1+\tau)^2\gamma^2_{_J}\gamma^2_{_R}\sigma^4_j - 4\gamma_{_J}\gamma_{_R}\sigma^2_j}}{2}.$$

Since $\gamma_{_J}\gamma_{_R}\sigma^2_j \leq \gamma_{_J}\gamma_{_R}\|L\|^2 < 1$, then ρ_i is complex and

$$|\rho_i| = \frac{1}{2}\sqrt{\big(2-(1+\tau)\gamma_{_J}\gamma_{_R}\sigma^2_j\big)^2 - \big((1+\tau)^2\gamma^2_{_J}\gamma^2_{_R}\sigma^4_j - 4\gamma_{_J}\gamma_{_R}\sigma^2_j\big)} = \sqrt{1-\tau\gamma_{_J}\gamma_{_R}\sigma^2_j} < 1.$$

As a result, $\lim_{k\to+\infty} D_i^k = 0$, M_{PD}^k is convergent and

$$M_{\mathrm{PD}}^\infty = \begin{bmatrix} Y & \\ & X \end{bmatrix} \begin{bmatrix} 0_l & & & \\ & \mathrm{Id}_{n-l} & & \\ & & 0_l & \\ & & & \mathrm{Id}_{m-l} \end{bmatrix} \begin{bmatrix} Y^T & \\ & X^T \end{bmatrix}$$

owing to (B.22), which is symmetric and positive semi-definite.

Trajectory of Primal–Dual With the above spectral properties of M_{PD}, we obtain immediately the trajectory of Primal–Dual owing to Proposition A.11. For the case $L = \mathrm{Id}$, then we have $\bar{L} = \mathcal{P}_{T_{w^\star}^{J^*}} \mathcal{P}_{T_{x^\star}^{R}}$, and the SVD values of $\bar{L}$ corresponds to the cosine value of the principal angle between $T_{x^\star}^R$ and $T_{w^\star}^{J^*}$... □

C Proofs of Section 5

For the easy of notation, let $M = M_{\mathscr{F}}$, $c = c_k, C = C_k$ and $\varepsilon = \varepsilon_k$.

Proof of theorem 6.2. Since $k \in \mathbb{N}$ is fixed throughout, we let $E_\ell \overset{\text{def}}{=} E_{k,\ell}$. We first show that for $\ell \in \mathbb{N}$, there holds

$$E_\ell = -\sum_{j=0}^{\ell-1} M^{\ell-1-j} F_{k+j} + \sum_{j=1}^{\ell} M^j E_0 C^{\ell-j} + \sum_{j=1}^{\ell} M^{\ell-j} F_{k-1} C^j. \tag{C.1}$$

We shall prove this by induction. First note that $E_0 = V_{k-1}C - V_k$

$$\begin{aligned} V_k C &\overset{(6.4)}{=} (MV_{k-1} + F_{k-1})C \overset{(6.5)}{=} MV_k + ME_0 + F_{k-1}C \\ &\overset{(6.4)}{=} V_{k+1} - F_k + ME_0 + F_{k-1}C. \end{aligned}$$

Hence, $E_1 = -F_k + ME_0 + F_{k-1}C$ and (C.1) is true for $\ell = 1$. Assume that (C.1) is true up to $\ell = m$, then,

$$\begin{aligned} V_k C^{m+1} \overset{(6.4)}{=} (MV_{k-1} + F_{k-1})C^{m+1} &= MV_{k-1}C^{m+1} + F_{k-1}C^{m+1} \\ &\overset{(6.5)}{=} MV_k C^m + ME_0 C^m + F_{k-1}C^{m+1} \\ &= M(V_{k+m} + E_m) + ME_0 C^m + F_{k-1}C^{m+1} \\ &\overset{(6.4)}{=} V_{k+m+1} - F_{k+m} + ME_m + ME_0 C^m + F_{k-1}C^{m+1}. \end{aligned}$$

Therefore, plugging in our assumption on E_m yields

$$\begin{aligned} E_{m+1} &= -F_{k+m} + ME_m + ME_0 C^m + F_{k-1}C^{m+1} \\ &= -F_{k+m} + \big(-\textstyle\sum_{j=0}^{m-1} M^{m-j} F_{k+j} + \sum_{j=1}^{m} M^{j+1} E_0 C^{m-j} + \sum_{j=1}^{m} M^{m+1-j} F_{k-1} C^j\big) + ME_0 C^m + F_{k-1}C^{m+1} \\ &= -\sum_{j=0}^{m} M^{m-j} F_{k+j} + \sum_{j=1}^{m+1} M^j E_0 C^{m+1-j} + \sum_{j=1}^{m+1} M^{m+1-j} F_{k-1} C^j. \end{aligned}$$

To bound the extrapolation error, observe that

$$\begin{aligned} \sum_{m=1}^{s} E_m &= \sum_{m=1}^{s} \big(-\textstyle\sum_{j=0}^{m-1} M^{m-1-j} F_{k+j} + \sum_{j=1}^{m} M^j E_0 C^{m-j} + \sum_{j=1}^{m} M^{m-j} F_{k-1} C^j\big) \\ &= -\sum_{\ell=0}^{s-1} M^\ell \sum_{j=0}^{s-1-\ell} F_{k+j} + \sum_{\ell=1}^{s} M^\ell E_0 \big(\textstyle\sum_{i=0}^{s-\ell} C^i\big) + \sum_{\ell=0}^{s-1} M^\ell F_{k-1} \big(\textstyle\sum_{i=1}^{s-\ell} C^i\big). \end{aligned} \tag{C.2}$$

Note that if $F_{k+j} = 0$ for all j, then

$$\sum_{m=1}^{s} E_m = \sum_{\ell=1}^{s} M^\ell E_0 \big(\textstyle\sum_{i=0}^{s-\ell} C^i\big)$$

and

$$\|\bar{z}_{k,s} - z^\star\| \leq \|z_{k+s} - z^\star\| + \sum_{\ell=1}^{s} \|M^\ell\| \|(\textstyle\sum_{i=0}^{s-\ell} C^i)_{(1,1)}\| \varepsilon.$$

In the general setting where $F_{k+j} \neq 0$, to bound the first term of (C.2), define

$$Z_k \stackrel{\text{def}}{=} \left[z_k | \cdots | z_{k-q+1}\right]$$

and note that $V_k = Z_k - Z_{k-1}$. So,

$$\sum_{j=0}^{m} F_{k+j} = \sum_{j=0}^{m} V_{k+j+1} - MV_{k+j} = Z_{k+m+1} - Z_k - MZ_{k+m} + MZ_{k-1},$$

and

$$\begin{aligned}
\big(\textstyle\sum_{\ell=0}^{s-1} M^\ell \sum_{j=0}^{s-1-\ell} F_{k+j}\big)_{(:,1)} &= \big(\textstyle\sum_{\ell=0}^{s-1} M^\ell Z_{k+s-\ell} - M^\ell Z_k - M^{\ell+1} Z_{k+s-1-\ell} + M^{\ell+1} Z_{k-1}\big)_{(:,1)} \\
&= \big(Z_{s+k} - M^s Z_k + \textstyle\sum_{\ell=0}^{s-1} M^\ell (MZ_{k-1} - Z_k)\big)_{(:,1)} \\
&= z_{k+s} - M^s z_k + \sum_{\ell=0}^{s-1} M^\ell (Mz_{k-1} - z_k) \\
&= z_{k+s} - M^s z_k + \sum_{\ell=0}^{s-1} M^\ell \big(M(z_{k-1} - z^\star) - (z_k - z^\star)\big) + \sum_{\ell=0}^{s-1} M^\ell (Mz^\star - z^\star) \\
&= z_{k+s} - M^s z_k + \sum_{\ell=0}^{s-1} M^\ell \big(M(z_{k-1} - z^\star) - (z_k - z^\star)\big) + M^s z^\star - z^\star \\
&= z_{k+s} - z^\star - M^s(z_k - z^\star) + \sum_{\ell=0}^{s-1} M^\ell \big(M(z_{k-1} - z^\star) - (z_k - z^\star)\big).
\end{aligned}$$

Therefore we arrive at,

$$\begin{aligned}
\|\bar{z}_{k,s} - z^\star\| \leq \|M^s(z_k - z^\star)\| &+ \|\textstyle\sum_{\ell=0}^{s-1} M^\ell\| \|(M(z_{k-1} - z^\star) - (z_k - z^\star)\| \\
&+ \sum_{\ell=1}^{s} \|M^\ell\| \|E_0 \textstyle\sum_{i=0}^{s-\ell} C^i_{(:,1)}\| + \sum_{\ell=0}^{s-1} \|M^\ell\| \|F_{k-1} \textstyle\sum_{i=1}^{s-\ell} C^i_{(:,1)}\|.
\end{aligned}$$

In the case of $s = +\infty$, we have

$$\begin{aligned}
\|\bar{z}_{k,\infty} - z^\star\| \leq \|(\mathrm{Id} - M)^{-1}\| &\|(M(z_{k-1} - z^\star) - (z_k - z^\star)\| \\
&+ \sum_{\ell=1}^{\infty} \|M^\ell\| \|E_0 (\mathrm{Id} - C)^{-1}_{(:,1)}\| + \sum_{\ell=0}^{\infty} \|M^\ell\| \|F_{k-1}((\mathrm{Id} - C)^{-1} - \mathrm{Id})_{(:,1)}\|.
\end{aligned}$$

We have $\|E_0(\mathrm{Id} - C)^{-1}_{(:,1)}\| = \frac{\varepsilon}{1 - \sum_i c_i}$ where

$$\varepsilon \stackrel{\text{def}}{=} \min_{c \in \mathbb{R}^q} \|\textstyle\sum_{j=1}^{q} c_j v_{k-j} - v_k\|.$$

Letting $b_i \stackrel{\text{def}}{=} \sum_{\ell=i}^{q} c_\ell$ for $i \geq 2$ and $b_1 = 1$,

$$F_{k-1}((\mathrm{Id} - C)^{-1} - \mathrm{Id})_{(:,1)} = \frac{1}{1 - \sum_\ell c_\ell} \sum_{i=1}^{q} b_i f_{k-i} - f_{k-1} = \frac{1}{1 - \sum_\ell c_\ell} \sum_{i=1}^{q} \big(\textstyle\sum_{\ell=i}^{q} c_\ell\big) f_{k-i}. \qquad \square$$

Proof of Theorem 6.4. The first result of Theorem 6.4 is simply a consequence of Theorem 6.2. To control the coefficients fitting error ε_k, we follow closely the arguments of [75, Section 6.7], since this amounts to understanding the behavior of the coefficients c_k, which are precisely the MPE coefficients. Recall our assumption that M is diagonalizable, so $M = U^\top \Sigma U$ where U is an orthogonal matrix and Σ is a diagonal matrix with the eigenvalues of M as its diagonal. Then, letting $u_k \overset{\text{def}}{=} Uv_k$,

$$\varepsilon_k = \min_{c\in\mathbb{R}^q} \|\sum_{i=1}^q c_i v_{k-i} - v_k\| = \min_{c\in\mathbb{R}^q} \|\sum_{i=1}^q c_i \Sigma^{k-i} u_0 - \Sigma^k u_0\| = \min_{g\in\mathscr{P}_q} \|\Sigma^{k-q} g(\Sigma) u_0\| \le \|u_0\| \min_{g\in\mathscr{P}_q} \max_{z\in\lambda(M)} |z|^{k-q}|g(z)|$$

where $\mathscr{P}_q$ is the set of monic polynomials of degree q and $\lambda(M)$ is the spectrum of M. Choosing $g = \prod_{j=1}^q (z-\lambda_j)$, we have $g(\lambda_j) = 0$ for $j = 1,\ldots,q$, so

$$\varepsilon_k \le \|u_0\| |\lambda_{q+1}|^{k-q} \max_{\ell>q} \prod_{j=1}^q |\lambda_j - \lambda_\ell|. \tag{C.3}$$

The claim that $\rho(C_k) < 1$ holds since the eigenvalues of C are precisely the roots of the polynomial $Q(z) = z^{k-1} - \sum_{i=1}^{k-1} c_j z^{k-1-i}$, and from [75], if $|\lambda_q| > |\lambda_{q+1}|$, then Q has precisely q roots $r_1,\ldots,r_q$ satisfying $r_j = \lambda_j + \mathscr{O}(|\lambda_{q+1}/\lambda_j|^k)$. So, $|r_j| < 1$ for all k sufficiently large. To prove the non-asymptotic bounds on ε_k, first observe that $z_{k+1} - z_k = M(z_k - z_{k-1})$ implies $z_{k+1} - z^\star = M(z_k - z_*)$ and $z_{k+1} - z_k = (M - \mathrm{Id})(z_k - z^\star)$. So, letting $\gamma_i = -c_{k,i}/(1-\sum_i c_{k,i})$ for $i = 1,\ldots,q$ and $\gamma_0 = 1/(1-\sum_i c_{k,i})$, we have

$$\frac{1}{1-\sum_i c_{k,i}}\left(v_k - \sum_{i=1}^q c_{k,i} v_{k-i}\right) = \sum_{i=0}^q \gamma_i v_{k-i} = (M-\mathrm{Id}) \sum_{i=0}^q \gamma_i (z_{k-i-1} - z^\star). \tag{C.4}$$

Now, $y \overset{\text{def}}{=} \sum_{i=0}^q \gamma_i z_{k-i-1}$ is precisely the MPE update and norm bounds on this are presented in [75]. For completeness, we reproduce their arguments here: Let $A \overset{\text{def}}{=} \mathrm{Id} - M$, by our assumption of $\lambda(M) \subset (-1,1)$, we have that A is positive definite. Then,

$$\|A^{1/2}(y-z^\star)\|^2 = \langle A(y-z^\star),\, (y-z^\star)\rangle = -\langle \sum_{i=0}^q \gamma_i v_{k-i},\, (y-z^\star)+w\rangle$$

where $w = \sum_{j=1}^q a_j v_{k-j}$ with $a \in \mathbb{R}^q$ being arbitrary, since by definition of γ, $\langle \sum_{i=0}^q \gamma_i v_{k-i},\, v_\ell\rangle = 0$ for all $\ell = k-q,\ldots,k-1$. We can write

$$w = \sum_{j=1}^q a_j (M-\mathrm{Id})(z_{k-j-1} - z^\star) = \sum_{j=1}^q a_j (M-\mathrm{Id}) M^{k-j-1}(z_0 - z^\star) = f(M)(z_{k-q-1} - z^\star)$$

where $f(z) = (z-1)\sum_{j=1}^q a_j z^{q-j}$, and we can write

$$y - z^\star = \sum_{i=0}^q \gamma_i M^{k-i-1}(z_0 - z^\star) = g(M)(z_{k-q-1} - z^\star)$$

where $g(z) = \sum_{i=0}^q \gamma_i z^{q-i}$. Therefore, $f(z) + g(z) = h(z)$, where h is a polynomial of degree q such that $h(1) = 1$. Moreover, since the coefficients a_j are arbitrary, h can be considered as an arbitrary element of $\tilde{\mathscr{P}}_q$, the set of all polynomials of degree q such that $h(1) = 1$. Therefore

$$\begin{aligned}\|A^{1/2}(y-z^\star)\|^2 &\le \|A^{1/2}(y-z^\star)\| \min_{h\in\tilde{\mathscr{P}}_q} \|h(M)(z_{k-q-1} - z^\star)\| \\ &\le \|A^{-1/2}(y-z^\star)\| \min_{h\in\tilde{\mathscr{P}}_q} \max_{t\in\lambda(M)} |h(t)| \|z_{k-q-1} - z^\star\|.\end{aligned}$$

In particular, combining this with (C.4), we have

$$\frac{\varepsilon_k}{|1-\sum_i c_{k,i}|} \le \|z_{k-q-1} - z^\star\| \|(\mathrm{Id}-M)^{1/2}\| \min_{h\in\tilde{\mathscr{P}}_q} \max_{t\in\lambda(M)} |h(t)|$$

Finally, in our case where $\lambda(M) = [\alpha,\beta]$ with $1 > \beta > \alpha > -1$, it is well known that $\min_{h\in\tilde{\mathscr{P}}_q} \max_{t\in\lambda(M)} |h(t)|$ has an explicit expression (see, for example, [13] or [75, Section 7.3.1]):

$$\min_{h\in\tilde{\mathscr{P}}_q} \max_{z\in\lambda(M)} |h(z)| \le \max_{z\in\lambda(M)} |h_*(z)|,$$

where $h_*(z) \overset{\text{def}}{=} \frac{T_q(\frac{2z-\alpha-\beta}{\beta-\alpha})}{T_q(\frac{2-\alpha-\beta}{\beta-\alpha})}$ where $T_q(x)$ is the q^{th} Chebyshev polynomial and it is well known that

$$\min_{h\in\tilde{\mathscr{P}}_q} \max_{z\in[\alpha,\beta]} |h(z)| \le 2\left(\frac{\sqrt{\eta}-1}{\sqrt{\eta}+1}\right)^q \tag{C.5}$$

where $\eta = \frac{1-\alpha}{1-\beta}$. □