

An Envelope for Davis-Yin Splitting and Strict Saddle Point Avoidance

Yanli Liu* Wotao Yin*

April 23, 2018

Abstract

It is known that operator splitting methods based on Forward Backward Splitting (FBS), Douglas-Rachford Splitting (DRS), and Davis-Yin Splitting (DYS) decompose a difficult optimization problems into simpler subproblems under proper convexity and smoothness assumptions. In this paper, we identify an envelope (an objective function) whose gradient descent iteration under a variable metric coincides with DYS iteration. This result generalizes the Moreau envelope for proximal-point iteration and the envelopes for FBS and DRS iterations identified by Patrinos, Stella, and Themelis.

Based on the new envelope and the Stable-Center Manifold Theorem, we further show that, when FBS or DRS iterations start from random points, they avoid all strict saddle points with probability one. This result extends the similar results by Lee et al. from gradient descent to splitting methods.

1 Introduction

The most general model considered in this paper minimizes the sum of three functions, where two of them are Lipschitz differentiable and, out of these two, one can involve a composition with a linear operator. The third function can be non-differentiable, and all the three functions can be nonconvex. A mathematical formulation is given in Section 3. The results of this paper, of course, apply to simpler models where any one or two of these three functions vanish. Problems that can be written in our general model are abundant. Examples include texture inpainting [16], matrix completion [6], and support vector machine classification [8].

Our model can be solved by the splitting iterative methods based on Douglas-Rachford Splitting (DRS) [15] and Forward-Backward Splitting (FBS) [17], as well as their generalization, Davis-Yin Splitting (DYS) [9]. In these methods,

*Department of Mathematics, University of California, Los Angeles, CA 90095.
E-mails: yanli / wotaoyin@math.ucla.edu

the problem objective is split into different steps, one of each of the objective functions. Their implementations are typically easy. By exploiting additional sum and coordinate friendly structures, they give rise to parallel and distributed algorithms that are highly scalable. The details of these methods are reviewed in Section 3 below.

These splitting methods are traditionally analyzed assuming that the sub-differentials of the objective functions are maximally monotone. Those of non-convex functions are generally non-monotone. Therefore, the majority of the existing results apply to only convex objective functions.

Recently, FBS and DRS are found to numerically converge for certain non-convex problems, for example, FBS for image restoration [22], dictionary learning, and matrix decomposition [25], and DRS for nonconvex feasibility problem [14], matrix completion [1], and phase retrieval [7]. Theoretically, their iterates have been shown to converge to stationary points in some nonconvex settings [2, 14, 26, 11]. In particular, any bounded sequence produced by FBS converges to a stationary point when the objective satisfies the KL property [2]; By using the Douglas-Rachford Envelope (DRE), DRS iterates are shown to converge to a stationary point when one of the two functions is Lipschitz differentiable, both of them are semi-algebraic and bounded below, and one of them is coercive [14, 26]; In [11], when one function is strongly convex and the other is weakly convex, and their sum is strongly convex, DRS iterates are shown to be Fejer monotone with respect to the set of fixed points of DRS operator, thus convergent. Though unlikely, it is still possible that the limit of a convergent sequence is a saddle point instead of a local minimum (except when all stationary points are local minima, which is the case studied in [11]). This depends on the problem geometry and the selected start point.

Recently, some first-order methods have been shown to avoid so-called strict saddle points, with probability one regarding random initialization [13, 12]. These results make a skillful use of the Stable-Center Manifold Theorem [20]. So far, their results apply to only relatively simple methods such as Gradient Descent, Coordinate Descent, and Proximal Point methods. We give a positive answer (under smoothness assumptions) that splitting methods also have this property. This result also matches the practical observations made in [24].

This paper makes the following contribution regarding the convergence and saddle point avoidance of FBS, DRS, and DYS iterations for nonconvex problems. We first generalize the existing Forward-Backward Envelope (FBE) and Douglas-Rachford Envelope (DRE) into a Davis-Yin Envelope (DYE) and establish relationships between the latter envelope and the original optimization objective. Then, under smoothness conditions, we show that the probability for DRS and FBS iterations with random initializations converge to strict saddle points of their respective DRE and FBE is zero. Finally, by the connection between the envelopes and the original objectives, we extend the above avoidance results to the strict saddle points of the original objectives. That is, when our problem has the strict saddle property, DRS and FBS with random initialization will almost surely converge to local minimizers. The strict saddle property is satisfied in several applications including, but not limited to, dic-

tionary learning [24], simple neural networks [5], phase retrieval [23], tensor decomposition [10], and low rank matrix factorization [4].

The rest of this paper is organized as follows. In Section 2, we introduce notation and review some useful results. In Section 3, we define the envelope for DYS and rewrite DYS equivalently as gradient descent of this envelope. In Section 4, we establish a strong relationship between the envelope and the objective. Then, in Section 5, we analyze the avoidance of strict saddle points of the objective. Finally, we conclude this paper in Section 6.

2 Preliminaries

In this section, we review some basic concepts, introduce our notation, and state some known results. For the sake of brevity, we omit proofs and direct references. We refer the reader to textbooks [19, 3].

We let $\mathbf{0} \in \mathbb{R}^n$ denote the vector zero, $\langle \cdot, \cdot \rangle$ the inner product, $\|\cdot\|$ the ℓ_2 norm, and $\text{Fix}T$ the set of fixed points of operator T .

A function $f: \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ is called β -weakly convex (or β -semiconvex) if the function $\tilde{f} := f + \frac{\beta}{2}\|\cdot\|^2$ is convex. Clearly, f can be nonconvex.

Let $y \xrightarrow{f} x$ denote $y \rightarrow x$ and $f(y) \rightarrow f(x)$. Then the subdifferential of f at $x \in \text{dom } f$ can be defined by

$$\partial f(x) := \left\{ v \in \mathbb{R}^n \mid \exists x^t \xrightarrow{f} x, v^t \rightarrow v, \text{ with } \liminf_{z \rightarrow x^t} \frac{f(z) - f(x^t) - \langle v^t, z - x^t \rangle}{\|z - x^t\|} \geq 0 \text{ for each } t \right\}.$$

If f is differentiable at x , we have $\partial f(x) = \{\nabla f(x)\}$; If f is convex, we have

$$\partial f(x) = \{v \in \mathbb{R}^n \mid f(z) \geq f(x) + \langle v, z - x \rangle \text{ for any } z \in \mathbb{R}^n\},$$

which is the classic definition of subdifferential in convex analysis.

A point x^* is a stationary point of a function f if $\mathbf{0} \in \partial f(x^*)$. x^* is a critical point of f if f is differentiable at x^* and $\nabla f(x^*) = \mathbf{0}$.

A point x^* is a *strict saddle point* of f if f is twice differentiable at x^* , x^* is a critical point of f , and $\lambda_{\min}[\nabla^2 f(x^*)] < 0$, where $\lambda_{\min}[\cdot]$ returns the smallest eigenvalue of the input. Local minimizers of a function are always its stationary points, but not strict saddle points.

For any $\gamma > 0$, the proximal mapping of a function f is defined by

$$P_{\gamma f}(x) := \arg \min_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{1}{2\gamma} \|y - x\|^2 \right\},$$

assuming that the right-hand side exists. When f is convex, $P_{\gamma f}$ is single-valued and satisfies $P_{\gamma f}(x) = (\text{Id} + \gamma \partial f)^{-1}(x)$, where Id is the identity map. For any proper, closed, convex function f , it Moreau Identity is

$$\text{Id} = P_{\gamma f} + \gamma P_{\frac{f^*}{\gamma}} \circ \frac{\text{Id}}{\gamma}, \quad (1)$$

where $f^*(u) := \sup_{x \in \mathbb{R}^n} \{\langle u, x \rangle - f(x)\}$ is the convex conjugate of f .

We also need the Inverse Function Theorem: for a C^1 mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, if the Jacobian $J_F(x)$ of F at $x \in \mathbb{R}^n$ is invertible, then there exists an inverse function F^{-1} defined in a neighbourhood of $F(x)$ such that F^{-1} is also C^1 and

$$J_{F^{-1}}(F(x)) = (J_F(x))^{-1}. \quad (2)$$

3 Envelope for Davis-Yin Splitting

In this section, we will introduce a function, which we call an envelope, such that DYS iteration can be written as gradient descent of this function under a variable metric. Since DYS generalizes FBS and DRS, the envelope of DYS is also a generalization of FBE and DRE, the respective envelopes of FBS and DRS, which were introduced in [21, 25].

3.1 Review of Davis-Yin splitting

DYS [9] can be applied to solve the following problem:

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \varphi(x) := f(x) + g(x) + h(Lx), \quad (3)$$

where, for this subsection, $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ and $h : \mathbb{R}^m \rightarrow \mathbb{R}$ are proper, closed, and convex, h is also L_h -Lipschitz differentiable, and $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear operator.

DYS iteration produces a sequence $(x^k)_{k \geq 0}$ according to $z^{k+1} = Tz^k$, where

$$Tz^k := z^k + \alpha \left(P_{\gamma f} \left(2P_{\gamma g}(z^k) - z^k - \gamma L^T \nabla h(Lz^k) \right) - P_{\gamma g}(z^k) \right),$$

where γ and α are positive scalars. We rewrite this operator into successive steps with designated letters as

$$\begin{aligned} q^k &:= L^T \nabla h(Lz^k), \\ r^k &:= 2P_{\gamma g}(z^k) - z^k, \\ p^k &:= P_{\gamma f}(r^k - \gamma q^k), \\ w^k &= p^k - P_{\gamma g}(z^k), \end{aligned} \quad (4)$$

$$z^{k+1} = Tz^k = z^k + \alpha w^k. \quad (5)$$

Since f, g are closed, proper, and convex (in this subsection), $P_{\gamma g}$ and $P_{\gamma f}$ are well defined and single valued. In [9], convergence is established for a range of parameters

$$\gamma \in \left(0, \frac{2L_h}{\|L\|^2} \right) \quad \text{and} \quad \alpha \in \left(0, 2 - \frac{\gamma \|L\|^2}{2L_h} \right).$$

When $h = 0$, (5) simplifies to Douglas-Rachford Splitting iteration,

$$z^{k+1} = z^k + \alpha(P_{\gamma f}(r^k) - P_{\gamma g}(z^k)).$$

When $g = 0$, $P_{\gamma g}$ reduces to Id and thus (5) simplifies to

$$z^{k+1} = z^k + \alpha(P_{\gamma f}(z^k - \gamma q^k) - z^k),$$

which is Forward-Backward Splitting iteration slightly generalized by including the linear operator L .

When $f = 0$, $P_{\gamma f}$ reduces to Id and (5) simplifies to Backward-Forward Splitting,

$$z^{k+1} = z^k + \alpha(P_{\gamma g}(z^k) - \gamma q^k - z^k).$$

When $f = g = 0$, (5) simplifies to gradient descent iteration

$$z^{k+1} = z^k - \alpha\gamma q^k.$$

3.2 Derivation of envelope

Now we show that, (5) can be written as gradient descent iteration of an envelope function under the following assumption.

Assumption 1.

1. $g : \mathbb{R}^n \rightarrow \mathbb{R}$ is L_g -Lipschitz differentiable.
2. $h : \mathbb{R}^m \rightarrow \mathbb{R}$ is L_h -Lipschitz differentiable.
3. $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ is proper, lower semicontinuous and $\gamma \in (0, \frac{1}{L_g + L_h \|L\|^2})$.
In addition, f is prox-bounded in the sense that $f(\cdot) + \frac{1}{2\gamma} \|\cdot\|^2$ is lower bounded for any $\gamma \in (0, \frac{1}{L_g + L_h \|L\|^2})$.

Compared to the assumption in Section 3.1, a main restriction is that g is Lipschitz differentiable. On the other hand, all f , g and h can be nonconvex.

We begin with two technical lemmas regarding the Moreau envelope of weakly convex functions and its twice differentiability.

Lemma 1. *Let ξ be proper, closed, β -weakly convex. Choose γ such that $\gamma \in (0, \frac{1}{\beta})$. Let $\xi^\gamma(z) = \min_{u \in \mathbb{R}^n} \{\xi(u) + \frac{1}{2\gamma} \|z - u\|^2\}$ be the Moreau envelope of ξ . Define $\tilde{\xi} = \xi + \frac{\beta}{2} \|\cdot\|^2$, which is convex. Then, proximal mapping $P_{\gamma\xi}(z)$ is single-valued and satisfies*

$$\begin{aligned} P_{\gamma\xi}(z) &= P_{\frac{\gamma}{1-\gamma\beta}\tilde{\xi}}(\frac{1}{1-\gamma\beta}z), \\ \nabla\xi^\gamma(z) &= \gamma^{-1}(z - P_{\gamma\xi}(z)). \end{aligned}$$

Furthermore, $P_{\gamma\xi}(z)$ is $\frac{1}{1-\gamma\beta}$ -Lipschitz continuous.

Proof. We have

$$\begin{aligned}\xi^\gamma(z) &= \min_{u \in \mathbb{R}} \left\{ \xi(u) + \frac{\beta}{2} \|u\|^2 + \frac{1}{2\gamma} \|u - z\|^2 - \frac{\beta}{2} \|u\|^2 \right\} \\ &= \min_{u \in \mathbb{R}} \left\{ \tilde{\xi}(u) + \frac{1 - \gamma\beta}{2\gamma} \|u - \frac{1}{1 - \gamma\beta} z\|^2 \right\} - \frac{\beta}{2 - 2\gamma\beta} \|z\|^2\end{aligned}$$

where the second equality follows from the definition of $\tilde{\xi}$.

As a result, for $\gamma \in (0, \frac{1}{\beta})$, $P_{\gamma\xi}$ is single-valued and

$$P_{\gamma\xi}(z) = P_{\frac{\gamma}{1 - \gamma\beta}\tilde{\xi}}\left(\frac{1}{1 - \gamma\beta}z\right).$$

Since $P_{\frac{\gamma}{1 - \gamma\beta}\tilde{\xi}}(z)$ is 1-Lipschitz continuous, we know that $P_{\gamma\xi}(z)$ is $\frac{1}{1 - \gamma\beta}$ -Lipschitz continuous.

Finally, since $\tilde{\xi}$ is convex, [3, Prop.12.29] tells us that ξ^γ is differentiable and

$$\begin{aligned}\nabla\xi^\gamma(z) &= \frac{1}{1 - \gamma\beta} \nabla\tilde{\xi}_{\frac{\gamma}{1 - \gamma\beta}}\left(\frac{1}{1 - \gamma\beta}z\right) - \frac{\beta}{1 - \gamma\beta}z \\ &= \frac{1}{1 - \gamma\beta} \frac{1 - \gamma\beta}{\gamma} \left(\frac{1}{1 - \gamma\beta}z - P_{\frac{\gamma}{1 - \gamma\beta}\tilde{\xi}}\left(\frac{1}{1 - \gamma\beta}z\right) \right) - \frac{\beta}{1 - \gamma\beta}z \\ &= \frac{1}{\gamma} (z - P_{\gamma\xi}(z)).\end{aligned}$$

□

Lemma 2. *In addition to Assumption 1, if g is twice differentiable at $P_{\gamma g}(z^0)$, then $P_{\gamma g}$ has a Jacobian at z^0 , g^γ is twice differentiable at z^0 with the Hessian*

$$\nabla^2 g^\gamma(z^0) = \frac{1}{\gamma} \left(I - \left(I + \gamma \nabla^2 g(P_{\gamma g}(z^0)) \right)^{-1} \right).$$

In addition, the mapping

$$A(z) := I - 2\gamma \nabla^2 g^\gamma(z) - \gamma L^T \nabla^2 h((LP_{\gamma g}(z))L(I - \gamma \nabla^2 g^\gamma(z))) \quad (6)$$

is invertible.

Proof. Since $\gamma \in (0, \frac{1}{L_g})$, $P_{\gamma g}$ is single-valued and

$$P_{\gamma g}(z^0) = (\text{Id} + \gamma \nabla g)^{-1} z^0, \quad (7)$$

where $(\text{Id} + \gamma \nabla g)^{-1}$ is the inverse mapping of $\text{Id} + \gamma \nabla g$. Since $\nabla^2 g(P_{\gamma g}(z^0))$ is symmetric and its eigenvalues are bounded by L_g , we know that $I + \gamma \nabla^2 g(P_{\gamma g}(z^0))$ is invertible, which is the Jacobian of $\text{Id} + \gamma \nabla g$ at $P_{\gamma g}(z^0)$.

Applying the Inverse Function Theorem to (7) by setting F as $P_{\gamma g}$ and z^0 as p in (2), we have

$$\nabla P_{\gamma g}(z^0) = \left(I + \gamma \nabla^2 g(P_{\gamma g}(z^0)) \right)^{-1},$$

Hence, Lemma 1 yields

$$\nabla^2 g^\gamma(z^0) = \frac{1}{\gamma} \left(I - \left(I + \gamma \nabla^2 g(P_{\gamma g}(z^0)) \right)^{-1} \right).$$

According to (6),

$$A(z^0) = A_1 - \gamma A_2. \quad (8)$$

where

$$\begin{aligned} A_1 &= 2 \left(I + \gamma \nabla^2 g(P_{\gamma g}(z^0)) \right)^{-1} - I, \\ A_2 &= L^T \nabla^2 h(L P_{\gamma g}(z^0)) L \left(I + \gamma \nabla^2 g(P_{\gamma g}(z^0)) \right)^{-1}. \end{aligned}$$

Since $\gamma \in (0, \frac{1}{L_g})$, A_1 is invertible, as a result,

$$\det(A(z^0)) = \det(A_1 - \gamma A_2) = \det(I - \gamma A_2 A_1^{-1}) \det(A_1) = \prod_{i=1}^n (1 - \gamma \lambda_i(A_2 A_1^{-1})) \det(A_1),$$

where $\lambda_i(A_2 A_1^{-1})$, $i = 1, \dots, n$ are the eigenvalues of $A_2 A_1^{-1}$.

Let us set

$$C = I + \gamma \nabla^2 g(P_{\gamma g}(z^0)) = C^T \succ 0,$$

and rewrite $A_2 A_1^{-1}$ as

$$A_2 A_1^{-1} = L^T \nabla^2 h(P_{\gamma g}(z^0)) L C^{-1} (2C - I)^{-1} = L^T \nabla^2 h(P_{\gamma g}(z^0)) L (2I - C)^{-1}.$$

Note that $L^T \nabla^2 h(P_{\gamma g}(z^0)) L$ is symmetric and $(2I - C)^{-1}$ is symmetric, positive definite. Therefore, $\lambda_i(A_2 A_1^{-1}) \in \mathbb{R}$, and we can set $\lambda_1(A_2 A_1^{-1}) \geq \lambda_2(A_2 A_1^{-1}) \dots \geq \lambda_n(A_2 A_1^{-1})$.

In order to show $\det(A(z^0)) \neq 0$, it suffices to show that $1 - \gamma \lambda_1 > 0$ when $\gamma \in (0, \frac{1}{L_g + L_h \|L\|_2^2})$.

We have

$$\begin{aligned} \lambda_1(A_2 A_1^{-1}) &\stackrel{(a)}{\leq} \lambda_1 \left(L^T \nabla^2 h(P_{\gamma g}(z^0)) L \right) \cdot \lambda_1((2I - C)^{-1}) \\ &\leq \lambda_1 \left(L^T \nabla^2 h(P_{\gamma g}(z^0)) L \right) \cdot \frac{1}{2 - (1 + \gamma L_g)} \\ &= \| (L^T \nabla^2 h(P_{\gamma g}(z^0)) L) \|_2 \cdot \frac{1}{1 - \gamma L_g} \\ &\stackrel{(b)}{\leq} \|L\|_2^2 \| \nabla^2 h(P_{\gamma g}(z^0)) \|_2 \frac{1}{1 - \gamma L_g} \\ &\leq \|L\|_2^2 L_h \frac{1}{1 - \gamma L_g}, \end{aligned}$$

where (a) is by [27, Corollary 11], and (b) is by Cauchy-Schwartz. Since $\gamma \in (0, \frac{1}{L_g + L_h \|L\|^2})$, we have

$$1 - \gamma \lambda_1(A_2 A_1^{-1}) > 1 - \|L\|_2^2 L_h \frac{\gamma}{1 - \gamma L_g} > 0.$$

Therefore, $\det(A(z^0)) \neq 0$. \square

We can now write DYS iteration (5) as gradient descent of an envelope under the following additional assumption.

Assumption 2.

1. f is β_f -weakly convex and $\gamma \in (0, \frac{1}{\beta_f})$.
2. g, h are twice continuously differentiable.

Theorem 3. Under Assumptions 1 and 2, DYS iteration (5) can be written equivalently as

$$z^{k+1} = T(z^k) = z^k - \alpha \gamma A^{-1}(z^k) \nabla \varphi^\gamma(z^k), \quad (9)$$

where the metric and envelope are, respectively,

$$\begin{aligned} A(z) &:= I - 2\gamma \nabla^2 g^\gamma(z) - \gamma L^T \nabla^2 h(L P_{\gamma g}(z)) L (I - \gamma \nabla^2 g^\gamma(z)), \\ \varphi^\gamma(z) &:= g^\gamma(z) - \gamma \|\nabla g^\gamma(z)\|^2 - \gamma \langle q(z), \nabla g^\gamma(z) \rangle + h(L P_{\gamma g}(z)) \\ &\quad - \frac{\gamma}{2} \|q(z)\|^2 + f^\gamma(z - 2\gamma \nabla g^\gamma(z) - \gamma q(z)). \end{aligned} \quad (10)$$

Proof. In view of Lemma 1 and (4), we have

$$w^k = p(z^k) - P_{\gamma g}(z^k) = P_{\gamma f}(r^k - \gamma q^k) - P_{\gamma g}(z^k), \quad (11)$$

where

$$\begin{aligned} P_{\gamma f}(r^k - \gamma q^k) &= r^k - \gamma q^k - \gamma \nabla f^\gamma(r^k - \gamma q^k), \\ P_{\gamma g}(z^k) &= z^k - \gamma \nabla g^\gamma(z^k), \\ r^k &= 2P_{\gamma g}(z^k) - z^k = z^k - 2\gamma \nabla g^\gamma(z^k), \\ q^k &= q(z^k) = L^T \nabla h(L(z^k - \gamma \nabla g^\gamma(z^k))). \end{aligned}$$

By substitution,

$$w^k = -\gamma \nabla g^\gamma(z^k) - \gamma q(z^k) - \gamma \nabla f^\gamma(z^k - 2\gamma \nabla g^\gamma(z^k) - \gamma q(z^k)).$$

Let ∇_z denote taking gradient to z ; then

$$\nabla_z f^\gamma(z - 2\gamma \nabla g^\gamma(z) - \gamma q(z)) = A(z) \nabla f^\gamma(z - 2\gamma \nabla g^\gamma(z) - \gamma q(z)),$$

where $A(z)$ is given in (6). After some computation we can verify that

$$\begin{aligned} A(z^k)w^k &= -\gamma(\nabla_z g^\gamma(z^k) - \gamma \nabla_z \|\nabla g^\gamma(z^k)\|^2) - \gamma\left(-\gamma \nabla_z (\langle q(z), \nabla g^\gamma(z^k) \rangle)\right) \\ &\quad - \gamma \nabla_z h(LP_{\gamma g}(z)) - \gamma\left(-\frac{\gamma}{2} \nabla_z \|q(z)\|^2\right) - \gamma \nabla_z f^\gamma\left(z^k - 2\gamma \nabla g^\gamma(z^k) - \gamma q(z)\right) \\ &= -\gamma \nabla \varphi^\gamma(z^k). \end{aligned}$$

Since $A(z^k)$ is invertible, we can rewrite DYS iteration (5) as (9). \square

4 Properties of envelope

In this section, we show that the global minimizers, local minimizers, critical(stationary) points, and strict saddle points of the envelope φ^γ defined in (10) correspond *one on one* to those of the objective function φ in (3).

First, we show lower and upper bounds of the envelope, which generalize [25, Prop. 2.3], [21, Prop. 4.3], and [18, Prop. 1].

Lemma 4. *Under Assumption 1, the following three inequalities hold for any $z \in \mathbb{R}^n$:*

$$\varphi^\gamma(z) \leq \varphi(P_{\gamma g}(z)), \quad (12)$$

$$\varphi^\gamma(z) \geq \varphi(p(z)) + C_1(\gamma)\|p(z) - P_{\gamma g}(z)\|^2, \quad (13)$$

$$\varphi^\gamma(z) \leq \varphi(p(z)) + C_2(\gamma)\|p(z) - P_{\gamma g}(z)\|^2, \quad (14)$$

where $\varphi^\gamma(z)$ is defined in (10),

$$C_1(\gamma) := \frac{1 - \gamma L_h \|L\|^2 - \gamma L_g}{2\gamma} > 0,$$

$$C_2(\gamma) := \frac{1 + \gamma L_h \|L\|^2 + \gamma L_g}{2\gamma} > 0,$$

and $p(z)$ is any element of $P_{\gamma f}(2P_{\gamma g}(z) - z - \gamma q(z))$.

Proof of inequality (12). By applying Lemma 1 to g , $\varphi^\gamma(z)$ can be written as

$$\begin{aligned} \varphi^\gamma(z) &= \min_u \left\{ g(u) + \frac{1}{2\gamma} \|z - u\|^2 \right\} - \gamma \left\| \frac{1}{\gamma} (z - P_{\gamma g}(z)) \right\|^2 \\ &\quad - \gamma \left\langle q(z), \frac{1}{\gamma} (z - P_{\gamma g}(z)) \right\rangle \\ &\quad + h(LP_{\gamma g}(z)) - \frac{\gamma}{2} \|q(z)\|^2 \\ &\quad + \min_u \left\{ f(u) + \frac{1}{2\gamma} \left\| -z + 2P_{\gamma g}(z) - \gamma q(z) - u \right\|^2 \right\}. \quad (15) \end{aligned}$$

Taking $u = P_{\gamma g}(z)$ in the two minimums of (15), we have

$$\begin{aligned}
\varphi^\gamma(z) &\leq g(P_{\gamma g}(z)) + \frac{1}{2\gamma}\|z - P_{\gamma g}(z)\|^2 - \gamma\left\|\frac{1}{\gamma}(z - P_{\gamma g}(z))\right\|^2 \\
&\quad - \langle q(z), z - P_{\gamma g}(z) \rangle \\
&\quad + h(LP_{\gamma g}(z)) - \frac{\gamma}{2}\|q(z)\|^2 \\
&\quad + f(P_{\gamma g}(z)) + \frac{1}{2\gamma}\| -z + P_{\gamma g}(z) - \gamma q(z) \|^2 \\
&= \varphi(P_{\gamma g}(z)).
\end{aligned}$$

□

Proof of inequality (13). According to [19, Thm. 1.25], we know that $P_{\gamma f}(2P_{\gamma g}(z) - z - \gamma q(z)) \neq \emptyset$ for $\gamma \in (0, \frac{1}{L_g + L_h\|L\|^2})$.

By taking $u = P_{\gamma g}(z)$ in the first minimum of (15) and $u = p(z) \in P_{\gamma f}(2P_{\gamma g}(z) - z - \gamma q(z))$ in the second, we have

$$\begin{aligned}
\varphi^\gamma(z) &= g(P_{\gamma g}(z)) + \frac{1}{2\gamma}\|z - P_{\gamma g}(z)\|^2 \\
&\quad - \gamma\langle q(z), \frac{1}{\gamma}(z - P_{\gamma g}(z)) \rangle \\
&\quad + h(LP_{\gamma g}(z)) - \frac{\gamma}{2}\|q(z)\|^2 \\
&\quad + f(p(z)) + \frac{1}{2\gamma}\| -z + 2P_{\gamma g}(z) - \gamma q(z) - p(z) \|^2. \tag{16}
\end{aligned}$$

By making use of

$$h(y) \geq h(x) - \langle \nabla h(y), x - y \rangle - \frac{L_h}{2}\|x - y\|^2 \text{ for any } x, y \in \mathbb{R}^m,$$

we arrive at

$$\begin{aligned}
\varphi^\gamma(z) &\geq g(P_{\gamma g}(z)) - \frac{1}{2\gamma}\|z - P_{\gamma g}(z)\|^2 \\
&\quad - \langle q(z), z - P_{\gamma g}(z) \rangle \\
&\quad + h(LP_{\gamma g}(z)) - \langle \nabla h(LP_{\gamma g}(z)), L(p(z) - P_{\gamma g}(z)) \rangle \\
&\quad - \frac{L_h}{2}\|L(p(z) - P_{\gamma g}(z))\|^2 - \frac{\gamma}{2}\|q(z)\|^2 \\
&\quad + f(p(z)) + \frac{1}{2\gamma}\|2P_{\gamma g}(z) - z - \gamma q(z) - p(z)\|^2.
\end{aligned}$$

Next, by making use of $\|a + b + c\|^2 = \|a\|^2 + \|b\|^2 + \|c\|^2 + 2\langle a, b \rangle + 2\langle b, c \rangle + 2\langle a, c \rangle$ for

$$\begin{aligned}
a &= P_{\gamma g}(z) - p(z), \\
b &= P_{\gamma g}(z) - z, \\
c &= -\gamma q(z),
\end{aligned}$$

we obtain

$$\begin{aligned}\varphi^\gamma(z) &\geq g(\mathbf{P}_{\gamma g}(z)) + h(Lp(z)) - \frac{Lh}{2} \|L(p(z) - \mathbf{P}_{\gamma g}(z))\|^2 \\ &\quad + f(p(z)) + \frac{1}{2\gamma} \|\mathbf{P}_{\gamma g}(z) - p(z)\|^2 + \langle \mathbf{P}_{\gamma g}(z) - p(z), \frac{1}{\gamma} (\mathbf{P}_{\gamma g}(z) - z) \rangle.\end{aligned}$$

Finally, by substituting

$$\begin{aligned}\nabla g(\mathbf{P}_{\gamma g}(z)) &= -\frac{1}{\gamma} (\mathbf{P}_{\gamma g}(z) - z), \\ g(y) &\geq g(x) - \langle \nabla g(y), x - y \rangle - \frac{Lg}{2} \|x - y\|^2 \quad \text{for any } x, y \in \mathbb{R}^n,\end{aligned}$$

we arrive at (13). \square

Proof of inequality (14). Similarly to the proof above, we can also start from (16) and apply

$$\begin{aligned}h(y) &\leq h(x) - \langle \nabla h(y), x - y \rangle + \frac{Lg}{2} \|x - y\|^2 \quad \text{for any } x, y \in \mathbb{R}^m, \\ g(y) &\leq g(x) - \langle \nabla g(y), x - y \rangle + \frac{Lg}{2} \|x - y\|^2 \quad \text{for any } x, y \in \mathbb{R}^n,\end{aligned}$$

which gives (14). \square

Now we can establish the direct connections between the global and local minimizers of φ^γ and those of φ . These results generalize [25, Prop. 2.3] and [21, Thm. 4.4].

Theorem 5. *Under Assumption 1, we have*

1. $\inf_{x \in \mathbb{R}^n} \varphi(x) = \inf_{z \in \mathbb{R}^n} \varphi^\gamma(z)$,
2. $\arg \min_{x \in \mathbb{R}^n} \varphi(x) = \mathbf{P}_{\gamma g} \left(\arg \min_{z \in \mathbb{R}^n} (\varphi^\gamma(z)) \right)$.

Proof of 1. From (12) we have

$$\inf_{z \in \mathbb{R}^n} \varphi^\gamma(z) \leq \inf_{x \in \mathbb{R}^n} \varphi(x),$$

If $\inf_{z \in \mathbb{R}^n} \varphi^\gamma(z) < \inf_{x \in \mathbb{R}^n} \varphi(x)$, then there exists $z_1 \in \mathbb{R}^n$ such that $\varphi^\gamma(z_1) < \inf_{x \in \mathbb{R}^n} \varphi(x)$. Then (13) gives

$$\inf_{x \in \mathbb{R}^n} \varphi(x) > \varphi^\gamma(z_1) \geq \varphi(p(z_1)) + C(\gamma) \|\mathbf{P}_{\gamma g}(z_1) - p(z_1)\|^2 \geq \varphi(p(z_1)),$$

which is a contradiction. \square

Proof of 2. Let us first show that

$$\arg \min_{x \in \mathbb{R}^n} \varphi(x) \subseteq P_{\gamma g} \left(\arg \min_{z \in \mathbb{R}^n} (\varphi^\gamma(z)) \right).$$

There is nothing to show if $\arg \min_{x \in \mathbb{R}^n} \varphi(x) = \emptyset$; If $\arg \min_{x \in \mathbb{R}^n} \varphi(x) \neq \emptyset$, then, for any $x^* \in \arg \min_{x \in \mathbb{R}^n} \varphi(x)$, we have $x^* = P_{\gamma g}(z^*)$ for $z^* = (I + \gamma \nabla g)(x^*)$. As a result, (12) and (13) give us

$$\inf_{x \in \mathbb{R}^n} \varphi(x) = \varphi(x^*) = \varphi(P_{\gamma g}(z^*)) \geq \varphi^\gamma(z^*) \geq \varphi(p(z^*)) + C(\gamma) \|P_{\gamma g}(z^*) - p(z^*)\|^2.$$

Which enforces $P_{\gamma g}(z^*) = p(z^*)$ and $\varphi(P_{\gamma g}(z^*)) = \varphi^\gamma(z^*)$. So for any $z \in \mathbb{R}^n$ we have

$$\varphi^\gamma(z^*) = \inf_{x \in \mathbb{R}^n} \varphi(x) \leq \varphi(p(z)) \leq \varphi^\gamma(z) - C(\gamma) \|P_{\gamma g}(z) - p(z)\|^2 \leq \varphi^\gamma(z),$$

which yields $z^* \in \arg \min_{z \in \mathbb{R}^n} \varphi^\gamma(z)$, $x^* \in P_{\gamma g} \left(\arg \min_{z \in \mathbb{R}^n} (\varphi^\gamma(z)) \right)$.

Now let us show that

$$P_{\gamma g} \left(\arg \min_{z \in \mathbb{R}^n} (\varphi^\gamma(z)) \right) \subseteq \arg \min_{x \in \mathbb{R}^n} \varphi(x).$$

Again, we can assume that $\arg \min_{z \in \mathbb{R}^n} (\varphi^\gamma(z)) \neq \emptyset$. For any $z^* \in \arg \min_{z \in \mathbb{R}^n} \varphi^\gamma(z)$, we need to show $P_{\gamma g}(z^*) \in \arg \min_{x \in \mathbb{R}^n} \varphi(x)$.

Let $z^{**} = (I + \gamma \nabla g)p(z^*)$, then $P_{\gamma g}(z^{**}) = p(z^*)$ and (12) and (13) give us

$$\varphi^\gamma(z^{**}) \leq \varphi(P_{\gamma g}(z^{**})) = \varphi(p(z^*)) \leq \varphi^\gamma(z^*) - C(\gamma) \|P_{\gamma g}(z^*) - p(z^*)\|^2.$$

Since $z^* \in \arg \min_{z \in \mathbb{R}^n} \varphi^\gamma(z)$, we must have

$$\begin{aligned} P_{\gamma g}(z^*) &= p(z^*) = P_{\gamma g}(z^{**}), \\ \varphi^\gamma(z^*) &= \varphi^\gamma(z^{**}) = \varphi(P_{\gamma g}(z^*)). \end{aligned}$$

Consequently, for any $z \in \mathbb{R}^n$ we have

$$\varphi(P_{\gamma g}(z^*)) = \varphi^\gamma(z^*) \leq \varphi^\gamma(z) \leq \varphi(P_{\gamma g}(z)),$$

which concludes $P_{\gamma g}(z^*) \in \arg \min_{x \in \mathbb{R}^n} \varphi(x)$. \square

It turns out that the local minimizers of φ and φ^γ also have one-to-one correspondence.

Theorem 6. *Under Assumptions 1 and 2, we have:*

1. *If $P_{\gamma g}(z^*) \in \arg \min_{x \in B(P_{\gamma g}(z^*), \delta)} \varphi(x)$ for some $\delta > 0$, then z^* is a local minimizer of φ^γ .*
2. *If $z^* \in \arg \min_{z \in B(z^*, \varepsilon)} \varphi^\gamma(z)$ for some $\varepsilon > 0$, then*

$$\varphi(P_{\gamma g}(z^*)) \leq \varphi(P_{\gamma g}(z)) \text{ for all } z \text{ such that } \|z - z^*\| \leq \varepsilon.$$

That is, $P_{\gamma g}(z^)$ is a local minimizer of $\varphi(x)$.*

Proof of 1. Since $P_{\gamma g}(z^*)$ is a local minimizer of φ , according to [19, Exercise 10.10], we have

$$\mathbf{0} \in \partial\varphi(P_{\gamma g}(z^*)) = \partial f(P_{\gamma g}(z^*)) + \nabla g(P_{\gamma g}(z^*)) + q(z^*).$$

Since $P_{\gamma g}$ is single-valued, this is equivalent to

$$\mathbf{0} \in \partial f(P_{\gamma g}(z^*)) + \frac{1}{\gamma}(-P_{\gamma g}(z^*) + z^* + \gamma q(z^*)),$$

Since $f + \frac{1}{2\gamma}\|\cdot\|^2$ is convex and $P_{\gamma f}$ is single valued, this is further equivalent to

$$P_{\gamma g}(z^*) = P_{\gamma f}(2P_{\gamma g}(z^*) - z^* - \gamma q(z^*)) = p(z^*).$$

According to Lemma 1, $P_{\gamma f}$ is $\frac{1}{1-\gamma\beta_f}$ -Lipschitz continuous, we can conclude that there exists $\eta > 0$ such that when $\|z - z^*\| \leq \eta$, we have $\|p(z) - p(z^*)\| \leq \delta$ and

$$\varphi^\gamma(z^*) = \varphi(P_{\gamma g}(z^*)) = \varphi(p(z^*)) \leq \varphi(p(z)) \leq \varphi^\gamma(z) - C(\gamma)\|P_{\gamma g}(z) - p(z)\|^2 \leq \varphi^\gamma(z).$$

□

Proof of 2. According to Lemma 2, $A(z)$ is invertible at z^* . Theorem 3 tells us that φ^γ is differentiable at z^* , so $\nabla\varphi^\gamma(z^*) = \mathbf{0}$ and $P_{\gamma g}(z^*) = p(z^*)$. As a result, for any $z \in \mathbb{R}^n$ with $\|z - z^*\| \leq \varepsilon$ we have

$$\varphi(P_{\gamma g}(z^*)) = \varphi^\gamma(z^*) \leq \varphi^\gamma(z) \leq \varphi(P_{\gamma g}(z)).$$

Furthermore, according to Lemma 2 we have

$$P_{\gamma g}(z) = P_{\gamma g}(z^*) + \left(I + \gamma\nabla^2 g(P_{\gamma g}(z^*))\right)^{-1}(z - z^*) + o(\|z - z^*\|).$$

Since $\left(I + \gamma\nabla^2 g(P_{\gamma g}(z^*))\right)^{-1}$ is positive definite, we can conclude that $P_{\gamma g}(B(z^*, \varepsilon))$ contains a ball centered at $P_{\gamma g}(z^*)$, as a result, $P_{\gamma g}(z^*)$ is a local minimizer of $\varphi(x)$. □

Now, let us show the one-to-one correspondence between the critical points of the envelope φ^γ and the stationary points of the objective $\varphi(x)$.

Theorem 7. *Under Assumptions 1 and 2, z^* is a critical point of φ^γ if and only if $P_{\gamma g}(z^*)$ is a stationary point of φ .*

Proof. Since f is β_f -weakly convex and $\gamma \in (0, \frac{1}{\beta_f})$, by Lemma 1, we know that $P_{\gamma f}$ is single-valued. And by Theorem 3, we have

$$\nabla\varphi^\gamma(z) = -A(z)\frac{1}{\gamma}(p(z) - P_{\gamma g}(z)), \quad (17)$$

where $p(z) = P_{\gamma f} \left((2P_{\gamma g}(z) - z - \gamma L^T \nabla h(LP_{\gamma g}(z))) \right)$. So $\nabla \varphi^\gamma(z^*) = \mathbf{0}$ if and only if

$$\begin{aligned} P_{\gamma g}(z^*) &= P_{\gamma f} \left(2P_{\gamma g}(z^*) - z^* - \gamma L^T \nabla h(LP_{\gamma g}(z^*)) \right) \\ &= \arg \min_z \left\{ f(z) + \frac{1}{2\gamma} \|z - (2P_{\gamma g}(z^*) - z^* - \gamma L^T \nabla h(LP_{\gamma g}(z^*)))\|^2 \right\}. \end{aligned}$$

Since the objective in the argmin is convex, by [19, Exercise 10.10] we know that this is equivalent to

$$\mathbf{0} \in \partial f(P_{\gamma g}(z^*)) + \frac{1}{\gamma} \left(-P_{\gamma g}(z^*) + z^* + \gamma L^T \nabla h(LP_{\gamma g}(z^*)) \right).$$

By the definition of $P_{\gamma g}$ and $\gamma \in (0, \frac{1}{L_g + L_h \|L\|^2})$, this is further equivalent to

$$\mathbf{0} \in \partial f(P_{\gamma g}(z^*)) + \nabla g(P_{\gamma g}(z^*)) + L^T \nabla h(LP_{\gamma g}(z^*)) = \partial \varphi(P_{\gamma g}(z^*)).$$

□

In order to establish the correspondence between the strict saddles of φ^γ and φ , we also need the following assumption.

Assumption 3. For any critical point z^* of φ^γ , f is twice continuously differentiable in a small neighbourhood of $P_{\gamma g}(z^*)$.

Lemma 8. Let z^* be a critical point of φ^γ . Under Assumptions 1, 2 and 3, $\varphi^\gamma(z)$ is twice differentiable at z^* and

$$\nabla^2 \varphi^\gamma(z^*) = -A(z^*) \frac{1}{\gamma} (J_p(z^*) - J_{P_{\gamma g}}(z^*)). \quad (18)$$

Proof. This follows from (17), $p(z^*) = P_{\gamma g}(z^*)$, and [21, Prop. 2.A.2]. □

Theorem 9. Let z^* be a critical point of $\varphi^\gamma(z)$. Under Assumptions 1, 2 and 3, z^* is a strict saddle point of $\varphi^\gamma(z)$ if and only if $P_{\gamma g}(z^*)$ is a strict saddle point of $\varphi(z)$.

Proof. According to Lemma 8, we know that $\nabla^2 \varphi^\gamma(z^*)$ exists.

Let z^* be a strict saddle point of $\varphi^\gamma(z)$, then Taylor expansion gives

$$\begin{aligned} \varphi^\gamma(z) &= \varphi^\gamma(z^*) + \frac{1}{2} (z - z^*)^T \nabla^2 \varphi^\gamma(z^*) (z - z^*) + o(\|z - z^*\|^2), \\ \varphi(p(z)) &= \varphi(p(z^*)) + \frac{1}{2} (p(z) - p(z^*))^T \nabla^2 \varphi(p(z^*)) (p(z) - p(z^*)) + o(\|p(z) - p(z^*)\|^2). \end{aligned}$$

On the other hand, (13) gives

$$\varphi^\gamma(z) \geq \varphi(p(z)).$$

Let $\nabla^2\varphi^\gamma(z^*)v = \lambda v$, where $\|v\| = 1$ and $\lambda < 0$. Setting $z - z^* = \alpha v$, we arrive at

$$\begin{aligned} & \varphi^\gamma(z^*) + \frac{1}{2}\lambda\alpha^2 + o(\alpha^2) \\ & \geq \varphi(p(z^*)) + \frac{1}{2}(p(z) - p(z^*))^T \nabla^2\varphi(p(z^*)) (p(z) - p(z^*)) + o(\|p(z) - p(z^*)\|^2). \end{aligned} \tag{19}$$

Furthermore, (12), (13) together with $p(z^*) = P_{\gamma g}(z^*)$ yield $\varphi^\gamma(z^*) = \varphi(p(z^*))$, combine this with (19) and $p(z) - p(z^*) = O(\|z - z^*\|) = O(\alpha)$, we conclude that $\lambda_{\min}(\nabla^2\varphi(P_{\gamma g}(z^*))) < 0$.

Similarly, let $P_{\gamma g}(z^*)$ be a strict saddle of $\varphi(z)$, then Taylor expansions gives

$$\begin{aligned} \varphi^\gamma(z) &= \varphi^\gamma(z^*) + \frac{1}{2}(z - z^*)^T \nabla^2\varphi^\gamma(z^*)(z - z^*) + o(\|z - z^*\|^2), \\ \varphi(P_{\gamma g}(z)) &= \varphi(P_{\gamma g}(z^*)) + \frac{1}{2}(P_{\gamma g}(z) - P_{\gamma g}(z^*))^T \nabla^2\varphi(P_{\gamma g}(z^*)) (P_{\gamma g}(z) - P_{\gamma g}(z^*)) \\ &\quad + o(\|P_{\gamma g}(z) - P_{\gamma g}(z^*)\|^2). \end{aligned}$$

On the other hand, (12) gives

$$\varphi^\gamma(z) \leq \varphi(P_{\gamma g}(z)),$$

Let $\nabla^2\varphi(P_{\gamma g}(z^*))v = \lambda v$ where $\lambda < 0$ and $\|v\| = 1$. By setting $z = (\text{Id} + \gamma\nabla g)(P_{\gamma g}(z^*) + \alpha v)$, we obtain $P_{\gamma g}(z) - P_{\gamma g}(z^*) = \alpha v$, taking $\alpha \rightarrow 0$ gives $\lambda_{\min}(\nabla^2\varphi^\gamma(z^*)) < 0$. \square

5 Avoidance of strict saddle points

In this section, we first show that under smoothness conditions, the probability for DRS and FBS with random initializations to converge to strict saddle points of DRE and FBE is zero, respectively. Then, by combining this result with the correspondence between the strict saddle points of the envelope and the objective, as stated in Theorem 9, we can conclude that DRS and FBS, if converge, will almost always avoid the strict saddle points of the objective. Therefore, when the objective satisfies the ‘‘strict saddle property’’, DRS and FBS, if they converge, will almost always converge to local minimizers.

In order to prove the main result, Theorem 14, we need the following Stable-Center Manifold Theorem, and its direct consequence, Theorem 11.

Theorem 10 states that, if T is a local diffeomorphism around one of its fixed point z^* , then there is a local stable center manifold $W_{\text{loc}}^{\text{cs}}$ with dimension equal to the number of eigenvalues of the Jacobian of T at z^* that are less than or equal to 1. Furthermore, there exists a neighbourhood B of z^* , such that a point z must be in $W_{\text{loc}}^{\text{cs}}$ if its forward iterations $T^k(z)$, for all $k \geq 0$, stay in B .

Theorem 10 (Theorem III.7, Shub [20]). *Let z^* be a fixed point for a C^r local diffeomorphism $T : U \rightarrow \mathbb{R}^n$, where U is a neighbourhood of z^* and $r \geq 1$. Suppose $E = E_s \oplus E_u$, where E_s is the span of the eigenvectors that correspond to eigenvalues of $J_T(z^*)$ that have magnitude less than, or equal to, 1, and E_u is the span of eigenvectors that correspond to eigenvalues of $J_T(z^*)$ that have magnitude greater than 1. Then there exists a C^r embedded disk $W_{\text{loc}}^{\text{cs}}$ that is tangent to E_s at z^* , which is called the local stable center manifold. Moreover, there exists a neighbourhood B of z^* , such that $T(W_{\text{loc}}^{\text{cs}}) \cap B \subseteq W_{\text{loc}}^{\text{cs}}$, $\bigcap_{k=0}^{\infty} T^{-k}(B) \subseteq W_{\text{loc}}^{\text{cs}}$, where $T^{-k}(B) = \{z \in \mathbb{R}^n \mid T^k(z) \in B\}$.*

The assumption of this following theorem is weaker than that of Theorem 2 of [12], as we do not assume that T is invertible in \mathbb{R}^n but only around A_T^* .

Theorem 11. *Assume that $T(z) = z + \alpha(p(z) - P_{\gamma g}(z))$ is a local diffeomorphism around $z^* \in A_T^*$, where $A_T^* = \{z \in \mathbb{R}^n \mid z = T(z), \max_i |\lambda_i(J_T(z))| > 1\}$ is the set of unstable fixed points of T . Then the set $W = \{z^0 : \lim z^k = z^*\}$ has Lebesgue measure $\mu(W) = 0$ in \mathbb{R}^n .*

Proof. Take $z^0 \in W$, so $z^k = T^k(z^0) \rightarrow z^* \in A_T^*$, there exists $T > 0$, such that for any $t \geq T$ we have $T^k(z^0) \in B_{z^*}$. So $T^t(z^0) \in S \triangleq \bigcap_{k=0}^{\infty} T^{-k}(B_{z^*})$ for any $t \geq T$.

From Theorem 10 we know that S is a subset of the local center stable manifold with codimension greater or equal to 1, so we have $\mu(S) = 0$;

Finally, $T^T(z^0) \in S$ implies that $z^0 \in T^{-T}(S) \subseteq \bigcup_{j=0}^{\infty} T^{-j}(S)$, since

$$T^{-j}(S) = T^{-j}\left(\bigcap_{k=0}^{\infty} T^{-k}(B_{z^*})\right) = \bigcap_{k=j}^{\infty} T^{-k}(B_{z^*}) \subseteq \bigcap_{k=0}^{\infty} T^{-k}(B_{z^*}) = S,$$

we can conclude that $\mu(W) = 0$. \square

Now we are ready to show the main result of this section, that when γ and α are small enough, it is impossible for DRS and FBS to converge to any strict saddle point of φ^γ , thus any strict saddle point of φ .

Lemma 12. *Under Assumptions 1 and 2, $z^* \in \text{Fix}T$ if and only if $\nabla\varphi^\gamma(z^*) = 0$.*

Proof. This follows directly from Theorem 3. \square

Theorem 13. *Define $Z^* = \{z^* \in \mathbb{R}^n \mid \nabla\varphi^\gamma(z^*) = 0, \lambda_{\min}(\nabla^2\varphi^\gamma(z^*)) < 0\}$ as the set of strict saddle points of φ^γ . Under Assumptions 1, 2, and 3, we set $\alpha \in (0, \alpha_0)$, where*

$$\alpha_0 = \frac{1}{\max\left(\lambda_{\max}(J_{P_{\gamma g}}(z^*) - J_P(z^*)), 0\right)},$$

then in the case of DRS ($h = 0$) and FBS ($g = 0$), the set $W = \{z^0 \in \mathbb{R}^n \mid \lim z^k = z^ \in Z^*\}$ satisfies $\mu(W) = 0$.*

Proof. Lemma 2 states that $A(z^*)$ is invertible. According to (18), for any $z^* \in X^*$, we have

$$\begin{aligned} \min_i \left(\lambda_i \left(-A^{\frac{1}{2}}(z^*) \frac{1}{\gamma} (J_p(z^*) - J_{P_{\gamma g}}(z^*)) A^{\frac{1}{2}}(z^*) \right) \right) &= \min_i \left(\lambda_i \left(A^{-\frac{1}{2}}(z^*) \nabla^2 \varphi^\gamma(z) A^{\frac{1}{2}}(z^*) \right) \right) \\ &= \min_i \left(\lambda_i \left(\nabla^2 \varphi^\gamma(z^*) \right) \right) < 0. \end{aligned}$$

Since $A(z)$ defined in (6) is symmetric when $h = 0$ or $g = 0$, we have

$$\max_i \left(\lambda_i \left(J_p(z^*) - J_{P_{\gamma g}}(z^*) \right) \right) > 0.$$

Since

$$\lambda_i(J_T(z^*)) = 1 + \alpha \lambda_i(J_p(z^*) - J_{P_{\gamma g}}(z^*)),$$

from Theorem 7, 9 and Lemma 12, we know that $Z^* \subseteq A_T^*$. Furthermore, since $J_p(z)$ is continuous around z^* and $\det(J_T(z^*)) \neq 0$ when $\alpha \in (0, \alpha_0)$, we know that T is a local diffeomorphism around z^* , therefore Theorem 11 gives $\mu(W) = 0$. \square

Theorem 14. Define $X^* = \{x^* \in \mathbb{R}^n \mid \nabla \varphi(x^*) = 0, \lambda_{\min}(\nabla^2 \varphi(x^*)) < 0\}$. Under Assumptions 1, 2, and 3, the set $V = \{z^0 \in \mathbb{R}^n \mid \lim P_{\gamma g}(z^k) = x^* \in X^*\}$ satisfies $\mu(V) = 0$.

Proof. Combine Theorem 9 with Theorem 13. \square

By Theorem 14, under smoothness conditions, DRS and FBS iterates will almost always avoid the strict saddle points of the objective. When the objective satisfies the *strict saddle property*, i.e., the saddle points of the objective are either local minimizers or strict saddle points, we can conclude that FBS and DRS almost always converge to local minimizers of the objective whenever they converge.

Many problems in practice satisfy the strict saddle property. Examples include dictionary learning [24], simple neural networks [5], phase retrieval [23], tensor decomposition [10], and low rank matrix factorization [4].

6 Conclusion

In this paper, we have constructed an envelope for DYS, and established various properties of this envelope. Specifically, there are one-to-one correspondences between the global, local minimizers, critical (stationary) points and strict saddle points of this envelope and those of the original objective. Then, by the Stable-Center Manifold theorem, we have shown that, the probability for FBS or DRS from random starting points to converge to strict saddle points of the envelope is zero. If the original objective also satisfies the strict saddle property, we have concluded that, whenever FBS and DRS converge, their iterates will almost always converge to local minimizers.

A limitation of this work lies in its smoothness assumptions. The construction of the envelope requires the Lipschitz differentiability of $g(x)$. Furthermore, twice differentiability of $f(x)$ at certain points is needed for the strict saddle avoidance property of FBS and DRS. It is certainly interesting to investigate the possibility of weakening these assumptions in the future.

Acknowledgements

This work is supported in part by NSF Grant DMS-1720237 and ONR Grant N000141712162.

References

- [1] Francisco J Aragón Artacho, Jonathan M Borwein, and Matthew K Tam. Douglas–rachford feasibility methods for matrix completion problems. *The ANZIAM Journal*, 55(4):299–326, 2014.
- [2] Hedy Attouch, Jérôme Bolte, and Benar Fux 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.
- [3] Heinz H Bauschke and Patrick L Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*, volume 408. Springer, 2011.
- [4] Srinadh Bhojanapalli, Behnam Neyshabur, and Nati Srebro. Global optimality of local search for low rank matrix recovery. In *Advances in Neural Information Processing Systems*, pages 3873–3881, 2016.
- [5] Alon Brutzkus and Amir Globerson. Globally optimal gradient descent for a convnet with gaussian inputs. In *International Conference on Machine Learning*, pages 605–614, 2017.
- [6] Emmanuel J Candes and Yaniv Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010.
- [7] Pengwen Chen and Albert Fannjiang. Fourier phase retrieval with a single mask by douglas–rachford algorithms. *Applied and Computational Harmonic Analysis*, 2016.
- [8] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [9] Damek Davis and Wotao Yin. A three-operator splitting scheme and its optimization applications. *Set-valued and variational analysis*, 25(4):829–858, 2017.

- [10] Rong Ge, Furong Huang, Chi Jin, and Yang Yuan. Escaping from saddle points—online stochastic gradient for tensor decomposition. In *Conference on Learning Theory*, pages 797–842, 2015.
- [11] Ke Guo, Deren Han, and Xiaoming Yuan. Convergence analysis of douglas–rachford splitting method for “strongly+ weakly” convex programming. *SIAM Journal on Numerical Analysis*, 55(4):1549–1577, 2017.
- [12] Jason D Lee, Ioannis Panageas, Georgios Piliouras, Max Simchowitz, Michael I Jordan, and Benjamin Recht. First-order methods almost always avoid saddle points. *arXiv preprint arXiv:1710.07406*, 2017.
- [13] Jason D Lee, Max Simchowitz, Michael I Jordan, and Benjamin Recht. Gradient descent only converges to minimizers. In *Conference on Learning Theory*, pages 1246–1257, 2016.
- [14] Guoyin Li and Ting Kei Pong. Douglas–rachford splitting for nonconvex optimization with application to nonconvex feasibility problems. *Mathematical programming*, 159(1-2):371–401, 2016.
- [15] Pierre-Louis Lions and Bertrand Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979.
- [16] Ji Liu, Przemyslaw Musialski, Peter Wonka, and Jieping Ye. Tensor completion for estimating missing values in visual data. *IEEE transactions on pattern analysis and machine intelligence*, 35(1):208–220, 2013.
- [17] Gregory B Passty. Ergodic convergence to a zero of the sum of monotone operators in hilbert space. *Journal of Mathematical Analysis and Applications*, 72(2):383–390, 1979.
- [18] Panagiotis Patrinos, Lorenzo Stella, and Alberto Bemporad. Douglas–rachford splitting: Complexity estimates and accelerated variants. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 4234–4239. IEEE, 2014.
- [19] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.
- [20] Michael Shub. *Global stability of dynamical systems*. Springer Science & Business Media, 2013.
- [21] Lorenzo Stella. Proximal envelopes: Smooth optimization algorithms for nonsmooth problems. 2017.
- [22] Lorenzo Stella, Andreas Themelis, and Panagiotis Patrinos. Forward–backward quasi-newton methods for nonsmooth optimization problems. *Computational Optimization and Applications*, 67(3):443–487, 2017.

- [23] Ju Sun, Qing Qu, and John Wright. A geometric analysis of phase retrieval. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 2379–2383. IEEE, 2016.
- [24] Ju Sun, Qing Qu, and John Wright. Complete dictionary recovery over the sphere i: Overview and the geometric picture. *IEEE Transactions on Information Theory*, 63(2):853–884, 2017.
- [25] Andreas Themelis, Lorenzo Stella, and Panagiotis Patrinos. Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone line-search algorithms. *arXiv preprint arXiv:1606.06256*, 2016.
- [26] Andreas Themelis, Lorenzo Stella, and Panagiotis Patrinos. Douglas-rachford splitting and admm for nonconvex optimization: new convergence results and accelerated versions. *arXiv preprint arXiv:1709.05747*, 2017.
- [27] Fuzhen Zhang and Qingling Zhang. Eigenvalue inequalities for matrix product. *IEEE Transactions on Automatic Control*, 51(9):1506–1509, 2006.