Unified Convergence Results on a Minimax Algorithm for Finding Multiple Critical Points in Banach Spaces

Xudong Yao[*] and Jianxin Zhou[†]

November 3, 2006

Abstract

A minimax method for finding multiple critical points in Banach spaces is successfully developed in [12] by using a projected pseudo-gradient as a search direction. Since several different techniques can be used to compute a projected pseudo-gradient, the uniform stepsize and the continuity of a search direction, two key properties for the convergence results in [5], get lost. In this paper, instead of proving convergence results of the algorithm for each technique, unified convergence results are obtained with a weaker stepsize assumption. An abstract existence-convergence result is also established. It is independent of the algorithm and explains why function values always converge faster than their gradients do. The weaker stepsize assumption is then verified for several different cases. As an illustration to the new results, the Banach space $W_0^{1,p}(\Omega)$ is considered and the conditions posed in the new results are verified for a quasilinear elliptic PDE.

[*]Department of Mathematics, University of Connecticut, Storrs, CT, 06269, xudong@math.uconn.edu.

[†]Department of Mathematics, Texas A&M University, College Station, TX, 77843, jzhou@math.tamu.edu. This author's research is supported in part by NSF DMS-0311905.

1   Introduction

Let B be a Banach space, B^* its topological dual, $\langle , \rangle$ the dual relation and $\| \cdot \|$ the norm in B. Let $J \in C^1(B, R)$ and $\nabla J \in B^*$ be its (Fréchet) gradient. A point $u^* \in B$ is a critical point of J if u^* solves the Euler-Lagrange equation $\nabla J(u^*) = 0$. The first candidates for a critical point are the local extrema. Traditional numerical algorithms focus on finding such *stable solutions*. Critical points that are not local extrema are *unstable* and called *saddle points*. In physical systems, saddle points appear as *unstable equilibria or transient excited states*. Multiple critical points exist in many nonlinear problems in applications ([3, 6, 7, 8, 10, 11], etc.). Choi-McKenna [1] in 1993 and Ding-Costa-Chen [2] in 1999 devised two algorithms for finding critical points of (the Morse index) MI=1 and MI=2, respectively. But no mathematical justification or convergence of the algorithms is established. Based on a local minimax characterization of saddle points, Li-Zhou [4] developed a local minimax algorithm (LMM) for finding critical points of MI=1,2,...,n, and proved its convergence in [5]. All those algorithms are formulated in Hilbert spaces where the gradient $\nabla J(u)$ played a key role to construct a search direction. In order to find multiple solutions in Banach spaces [3, 10], Yao-Zhou successfully developed the first LMM in Banach spaces and solved several quasilinear elliptic PDEs for multiple solutions [12]. The method is also modified to solve the nonlinear p-Laplacian operator for multiple eigen-pairs [13]. The key to the success of Yao-Zhou's algorithm is to replace the gradient by a projected pseudo-gradient (PPG). The purpose of this paper is to establish some convergence results for the algorithm.

Compare to those results in Hilbert spaces [5], there are several significant differences. When B is a Hilbert space, the gradient $\nabla J(u)$ which played the key role to construct a search direction in LMM in [5], is uniquely determined in B and naturally continuous if J is C^1 and $B = L \oplus L^{\perp}$ holds for any closed subspace L. When B is a Banach space, however, the gradient $\nabla J(u)$ is in B^* not B and cannot be directly used as a search direction in B. Thus a PPG is introduced to LMM. Although theoretically a Lipschitz continuous PPG flow exists, for most cases, no explicit formula is available. On the other hand, there are many different ways to select a PPG. When PPGs are numerically computed in an implementation, they may belong to different PPG flows. We lost the uniform stepsize property and the continuity of a search direction, two key conditions in proving the convergence results in [5]. To make

up the first loss, we design a weaker stepsize condition, assumption (H), to replace the old uniform stepsize property; To make up the second loss, we generalize the notion of a peak selection to that of an L-$\perp$ selection with which its continuity or smoothness can be verified. Thus corresponding modifications in LMM [12] have to be made.

To simplify our approach, in this paper, we assume $B = L \oplus L'$. When L is finite-dimensional, such L' always exists. In particular, for the commonly used Banach space $W_0^{1,p}(\Omega)$, we present an explicit formula for obtaining L' and a practical technique to compute a PPG.

Instead of proving convergence results of the algorithm for each of the techniques used to compute a PPG, in this paper, we establish some unified convergence results. To do so, in Section 2, we generalize a peak selection to an L-$\perp$ selection and prove the existence of a PPG at a value of an L-$\perp$ selection. A new LMM and its mathematical foundation are also presented there. Sections 3 is devoted to prove unified convergence results. We introduce a new stepsize assumption (H) and then prove a subsequence convergence result, Theorem 3.1, under some very reasonable assumptions. An abstract existence-convergence result, Theorem 3.2 is then established. This result is actually independent of algorithms. It also explains why in our LMM, function values always converge faster than their gradients do. Based on this abstract result, another convergence result, Corollary 3.1, is proved to show that under certain conditions, a convergent subsequence implies a point-to-set convergence. Assumption (H) is then verified for several different cases, in particular, for the commonly used Banach space $W_0^{1,p}(\Omega)$. In the last section, we discuss how to check other conditions we posed in the convergence results. In particular, we present a quasilinear elliptic PDE and verify those conditions.

2   A Min-$\perp$ Method

Let L be a closed subspace of B. For a subspace $A \subseteq B$, denote $S_A = \{v \in A | \|v\| = 1\}$. For a point $v \in S_B$, let $[L, v] = \{tv + w | w \in L, t \in \mathbb{R}\}$. Since $\nabla J(u) \in B^*$ not B, it cannot be used as a search direction in B. Thus a pseudo-gradient is used instead.

DEFINITION 2.1. *Let $J : B \to \mathbb{R}$ be Fréchet differentiable at $u \in B$ with $\nabla J(u) \neq 0$ and $0 < \theta \leq 1$ be given. A point $\Psi(u) \in B$ is a* pseudo-gradient *of J at u w.r.t. θ if*

$$(2.1) \qquad \|\Psi(u)\| \leq 1, \quad \langle \nabla J(u), \Psi(u) \rangle \geq \theta \|\nabla J(u)\|.$$

A pseudo-gradient flow *of J w.r.t. θ is a continuous mapping $\mathcal{F} : B \to B$ such that $\forall u \in B$ with $\nabla J(u) \neq 0$, $\mathcal{F}(u)$ is a pseudo-gradient of J at u w.r.t. θ.*

In Definition 2.1, the condition $\|\Psi(u)\| \leq 1$ is not essential. It can be replaced by any bound $M \geq 1$, since after a normalization, θ can always be replaced by $\frac{\theta}{M}$. It is known [9] that a C^1 functional has a locally Lipschitz continuous pseudo-gradient flow. Pseudo-gradients have been used in the literature to find a minimum of a functional in Banach spaces. However, as saddle points are concerned, such pseudo-gradients do not help much, since they lead to a local minimum point. Thus we introduce a new notion, called a projected pseudo-gradient (PPG), which plays a key role in the success of our LMM in Banach spaces.

DEFINITION 2.2. *An L'-projected pseudo-gradient (PPG) $G(u)$ of J is a pseudo-gradient of J at u such that $G(u) \in L'$.*

The motivation to define a PPG is two-fold. Firstly, as a pseudo-gradient, it provides a proper searching termination criterion, i.e., with (2.1), $G(u) = 0$ implies $\nabla J(u) = 0$; Secondly the condition $G(u) \in L'$ is to prevent a pseudo-gradient search from entering the subspace L, which is spanned by previously found critical points. The existence of such L'-PPG of J at $u = \mathcal{P}(v)$ is established by Lemma 2.1 in [12], where $\mathcal{P}$ is a peak selection defined below.

DEFINITION 2.3. *([12]) A set-valued mapping $P : S_{L'} \to 2^B$ is called the* peak mapping *of J w.r.t. L if*

$$P(v) = \{w = \arg \text{ local-} \max_{u \in [L,v]} J(u)\}, \quad \forall v \in S_{L'}.$$

A mapping $\mathcal{P} : S_{L'} \to B$ is called a peak selection *of J w.r.t. L if $\mathcal{P}(v) \in P(v), \forall v \in S_{L'}$. If a peak selection $\mathcal{P}$ is* locally *defined near a point $v \in S_{L'}$, we say that J has a* local peak selection *$\mathcal{P}$ at v.*

By using a peak selection and a PPG, a local minimax method (LMM) is successfully developed in [12] for computing multiple saddle points in Banach spaces. However, as convergence analysis is concerned, such an algorithm has an ill-condition, i.e., the graph defined by a peak selection is not closed, in other words, a limit of a sequence of local

maxima is not necessarily a local maximum point. Consequently, we cannot talk about a limit, continuity or do convergence analysis within the content of a peak selection. We introduce a generalized notion.

DEFINITION 2.4. *A set-valued mapping* $P : S_{L'} \to 2^B$ *is called the L-$\perp$ mapping of J if*

$$P(v) = \{u \in [L, v] : \langle \nabla J(u), w \rangle = 0, \ \forall w \in [L, v]\}, \quad \forall v \in S_{L'}.$$

A mapping $\mathcal{P} : S_{L'} \to B$ *is called an L-$\perp$ selection of J if* $\mathcal{P}(v) \in P(v), \forall v \in S_{L'}$. *If an L-$\perp$ selection* $\mathcal{P}$ *is locally defined near a given* $v \in S_{L'}$, *we say that J has a local L-$\perp$ selection* $\mathcal{P}$ *at* v.

LEMMA 2.1. *If J is* C^1, *then the graph* $Gr = \{(u, v) : v \in S_{L'}, u \in P(v) \neq \emptyset\}$ *is closed.*

Proof. Let $(u_n, v_n) \in Gr$ and $(u_n, v_n) \to (u_0, v_0)$. We have $u_n \in [L, v_n]$, $\nabla J(u_n) \perp [L, v_n]$. Since $v_n \to v_0 \in S_{L'}$, for each $w \in [L, v_0]$, there are $w_n \in [L, v_n]$ such that $w_n \to w$. Thus $\nabla J(u_n) \perp w_n$. But J is C^1, $u_n \to u_0$ and $w_n \to w$ lead to $\nabla J(u_0) \perp w$, i.e., $\nabla J(u_0) \perp [L, v_0]$. Thus $v_0 \in S_{L'}$ and $u_0 \in P(v_0)$, i.e., $(u_0, v_0) \in Gr$. ∎

It is clear that if $\mathcal{P}$ is a peak selection of J w.r.t. L, then $\mathcal{P}$ is an L-$\perp$ selection of J. The generalization not only removes the ill-condition and makes it possible to check the continuity of $\mathcal{P}$ but also exceeds the scope of a minimax principle, the most popular approach in critical point theory. It enables us to treat non-minimax type saddle points, such as the monkey saddles, or a problem without a mountain pass structure, see Example 2.1 in [14]. By a similar argument as in Lemma 2.1 of [12] we can prove the following existence of an L'-PPG.

LEMMA 2.2. *Assume* $B = L \oplus L'$ *and* $0 < \theta < 1$ *be given. For* $v_0 \in S_{L'}$, *if* $\mathcal{P}$ *is a local L-$\perp$ selection of J at* v_0 *s.t.* $\nabla J(\mathcal{P}(v_0)) \neq 0$ *and* $\Psi(\mathcal{P}(v_0)) \in B$ *is a pseudo-gradient of J at* $\mathcal{P}(v_0)$ *w.r.t.* θ, *then there exists an* L'-PPG $G(\mathcal{P}(v_0))$ *of J at* $\mathcal{P}(v_0)$ *w.r.t.* θ *s.t.*

(a) $G(\mathcal{P}(v_0)) \in L'$, $0 < \|G(\mathcal{P}(v_0))\| \leq M := \|\mathbb{P}\|$ *where* $\mathbb{P} : B \to L'$ *is the linear projection;*

(b) $\langle \nabla J(\mathcal{P}(v_0)), G(\mathcal{P}(v_0)) \rangle = \langle \nabla J(\mathcal{P}(v_0)), \Psi(\mathcal{P}(v_0)) \rangle$;

(c) If $\Psi(\mathcal{P}(v_0))$ *is the value of a pseudo-gradient flow* $\Psi(\cdot)$ *of J at* $\mathcal{P}(v_0)$, *then* $G(\cdot)$ *is continuous and* $G(\mathcal{P}(v_0))$ *is called the value of an* L'-PPG *flow of J at* $\mathcal{P}(v_0)$.

We now start to establish some mathematical foundation for our new algorithm.

LEMMA 2.3. *([12])* $\quad \|v - \dfrac{v - w}{\|v - w\|}\| \leq \dfrac{2\|w\|}{\|v - w\|}, \quad \forall v \in B, \|v\| = 1, \forall w \in B.$ ∎

LEMMA 2.4. *For $v_0 \in S_{L'}$, if J has a local L-$\perp$ selection $\mathcal{P}$ at v_0 satisfying (1) $\mathcal{P}$ is continuous at v_0, (2) $d(\mathcal{P}(v_0), L) > 0$ and (3) $\nabla J(\mathcal{P}(v_0)) \neq 0$. Then, there exists $s_0 > 0$ such that for $0 < s < s_0$*

(2.2) $$J(\mathcal{P}(v_0(s))) - J(\mathcal{P}(v_0)) < -\frac{\theta s}{4}|t_0|\|\nabla J(\mathcal{P}(v_0))\|$$

where $\mathcal{P}(v_0) = t_0 v_0 + w_0$ for some $t_0 \in \mathbb{R}, w_0 \in L$, $v_0(s) = \dfrac{v_0 - sign(t_0)sG(\mathcal{P}(v_0))}{\|v_0 - sign(t_0)sG(\mathcal{P}(v_0))\|}$ and $G(\mathcal{P}(v_0))$ is an L'-PPG of J w.r.t. θ at $\mathcal{P}(v_0)$. ∎

The proof of Lemma 2.4 can follow a similar argument of Lemma 2.4 in [12]. The inequality (2.2) will be used to define a stepsize rule for the algorithm and establish convergence results. With Lemma 2.4, the following characterization of saddle points is clear.

THEOREM 2.1. *Let $v_0 \in S_{L'}$. Assume that J has a local L-$\perp$ selection $\mathcal{P}$ at v_0 such that (1) $\mathcal{P}$ is continuous at v_0, (2) $d(\mathcal{P}(v_0), L) > 0$ and (3) v_0 is a local minimum point of $J(\mathcal{P}(v))$. Then $\mathcal{P}(v_0)$ is a critical point of J.*

2.1   A Min-Orthogonal Algorithm

DEFINITION 2.5. *Let $v_0 \in S_{L'}$ and $\mathcal{P}$ be a local L-$\perp$ selection of J at v_0 with $\nabla J(\mathcal{P}(v_0)) \neq 0$. A point $w \in L'$ is a descent direction of $J(\mathcal{P}(\cdot))$ at v_0 if there is $s_0 > 0$ such that*

$$J(\mathcal{P}(v_0(s))) < J(\mathcal{P}(v_0)), \quad \forall \, 0 < s < s_0 \quad where \quad v_0(s) = \frac{v_0 + sw}{\|v_0 + sw\|}.$$

By Theorem 2.1, a descent direction method to approximate a local minimum of $J(\mathcal{P}(v))$ leads to the following min-$\perp$ algorithm.

Assume that $L = [u^1, u^2, ..., u^{n-1}]$, where $u^1, u^2, ..., u^{n-1}$ are $n-1$ previously found critical points of J and L' is a subspace of B such that $B = L \oplus L'$. For given positive numbers $\lambda, \theta \in (0, 1)$ and ε.

Step 1: Let $v_1 \in S_{L'}$ be an ascent-descent direction at u^{n-1}.

Step 2: Set $k = 1$. Solve for $u_k \equiv \mathcal{P}(v_k) \equiv t_0^k v_k + t_1^k u^1 + \cdots + t_{n-1}^k u^{n-1}$ such that $t_0^k > 0$,

$$\langle \nabla J(\mathcal{P}(v_k)), v_k \rangle = 0 \quad and \quad \langle \nabla J(\mathcal{P}(v_k)), u^i \rangle = 0, \; i = 1, 2, ..., n-1.$$

Step 3: Find a descent direction $w_k \in L'$ of $J(\mathcal{P}(\cdot))$ at v_k.

Step 4: If $\|\nabla J(u_k)\| \leq \varepsilon$, then output $u_k = \mathcal{P}(v_k)$, stop. Otherwise, do Step 5.

Step 5: For each $s > 0$, denote $v_k(s) = \dfrac{v_k + sw_k}{\|v_k + sw_k\|}$ and set $v_{k+1} = v_k(s_k)$ where

$$s_k = \max_{m \in \mathbb{N}} \left\{ \frac{\lambda}{2^m} \Big| 2^m > \|w_k\|, J(\mathcal{P}(v_k(\frac{\lambda}{2^m}))) - J(u_k) < -\frac{\theta|t_0^k|}{4}(\frac{\lambda}{2^m})\|\nabla J(u_k)\| \right\}.$$

Step 6: Update $k = k + 1$ and go to Step 3.

REMARK 2.1.

(1) The constant $\lambda \in (0, 1)$ is used to prevent the stepsize from being too large to loose search stability. From now on we always assume that λ is such a constant.

(2) In Step 2, one way to solve the equations while satisfying the nondegenerate condition $t_0^k > 0$ is to find a local maximum point u_k of J in the subspace $[L, v_k]$, i.e., $u_k = \mathcal{P}(v_k)$ and $\mathcal{P}$ is a peak selection of J w.r.t. L.

(3) In Step 3, we may assume $\|w_k\| \leq M$ for some $M \geq 1$. There are many different ways to select a descent direction w_k. However, when a descent direction is selected, a corresponding stepsize rule in Step 5 has to be designed so that it can be achieved and leads to a convergence. For example, when a negative L'-PPG flow, or a negative L'-PPG is used as a descent direction, we have $v_k \in S_{L'}$ and a positive stepsize s_k for the current stepsize rule in Step 5 can always be obtained. In some cases, when $-\nabla J(\mathcal{P}(v_k))$ is used to construct a descent direction, the stepsize rule in Step 5 has to be modified as in Case 3 below.

3   Unified Convergence Results

DEFINITION 3.1. *For each $v \in S_{L'}$ with $\|\nabla J(\mathcal{P}(v))\| \neq 0$, write $\mathcal{P}(v) = t_v v + v_L$ for some $v_L \in L$ and define the stepsize $s(v)$ at v as*

$$s(v) = \max_{m \in \mathbb{N}} \left\{ s = \frac{\lambda}{2^m} \Big| 2^m > \|w\|, J(\mathcal{P}(v(s))) - J(\mathcal{P}(v)) < -\frac{1}{4}\theta|t_v|s\|\nabla J(\mathcal{P}(v))\| \right\}$$

where w is a descent direction J at $\mathcal{P}(v)$ and $v(s) = \dfrac{v + sw}{\|v + sw\|}$.

Let $\{u_k\}$ be the sequence generated by the algorithm where $u_k = \mathcal{P}(v_k)$. Since a PPG can be computed by many different ways, we lost the uniform stepsize, one of the key condition, in [5]. Here we design a new stepsize assumption (H) to establish unified convergence results. This condition is weaker than the uniform stepsize assumption in [5] and will be verified for several different cases.

Assumption (**H**) :   if $v_0 \in S_{L'}$ with $\nabla J(\mathcal{P}(v_0)) \neq 0$ and $v_k \to v_0$, then

there is $s_0 > 0$ such that $s_k = s(v_k) \geq s_0$ when k is large.

We need the following PS condition and Ekeland's variational principle [10].

DEFINITION 3.2. *A function $J \in C^1(B, \mathbb{R})$ is said to satisfy the Palais-Smale (PS) condition if any sequence $\{u_i\} \subset B$ such that $\{J(u_i)\}$ is bounded and $\nabla J(u_i) \to 0$ possesses a convergent subsequence.*

LEMMA 3.1. *(Ekeland's variational principle) Let X be a complete metric space and $J : X \to \mathbb{R} \cup \{+\infty\}$ be a lower semi-continuous function bounded from below. Then, for any $\varepsilon > 0$ and $x_0 \in X$ with $J(x_0) < +\infty$, there is $\bar{x} \in X$ such that*

$$J(\bar{x}) + \varepsilon d(x_0, \bar{x}) \leq J(x_0) \quad and \quad J(x) + \varepsilon d(x, \bar{x}) > J(\bar{x}), \ \forall x \in X \ and \ x \neq \bar{x}. \quad \blacksquare$$

First we prove a subsequence convergence result whose conditions will be verified with an application problem in Section 4.

THEOREM 3.1. *Let $J \in C^1(B, \mathbb{R})$ satisfy the PS condition. If an L-$\perp$ selection $\mathcal{P}$ of J satisfies (1) $\mathcal{P}$ is continuous on $S_{L'}$, (2) $d(\mathcal{P}(v_k), L) \geq \alpha > 0, \forall k = 1, 2, ...,$ (3) $\inf\limits_{1 \leq k < \infty} J(\mathcal{P}(v_k)) > -\infty$, (4) assumption (H) is satisfied, then*
(a) $\{v_k\}$ has a subsequence $\{v_{k_i}\}$ such that $u_{k_i} = \mathcal{P}(v_{k_i})$ converges to a critical point of J;
(b) if a subsequence $v_{k_i} \to v_0$ as $i \to \infty$, then $u_0 = \mathcal{P}(v_0)$ is a critical point of J.

Proof. (a) By the stepsize rule and Lemma 2.3, for $k = 1, 2, ...$, we have

$$(3.1) \ J(u_{k+1}) - J(u_k) \leq -\frac{1}{4}\theta\alpha s_k \|\nabla J(\mathcal{P}(v_k))\| \leq -\frac{|1-\lambda|}{4M}\theta\alpha\|v_{k+1} - v_k\|\|\nabla J(\mathcal{P}(v_k))\|.$$

Suppose that there is $\delta > 0$ such that $\|\nabla J(\mathcal{P}(v_k))\| \geq \delta$ for any k. From (3.1), we have

$$(3.2) \qquad J(u_{k+1}) - J(u_k) \leq -\frac{|1-\lambda|}{4M}\theta\alpha\delta\|v_{k+1} - v_k\|, \quad \forall k = 0, 1, 2, ...$$

Adding up two sides of (3.2) gives

$$(3.3) \qquad \lim_{k \to \infty} J(u_k) - J(u_0) = \sum_{k=0}^{\infty}[J(u_{k+1}) - J(u_k)] \leq -\frac{|1-\lambda|}{4M}\theta\alpha\delta\sum_{k=0}^{\infty}\|v_{k+1} - v_k\|,$$

i.e., $\{v_k\}$ is a Cauchy sequence. Thus $v_k \to \hat{v} \in S_{L'}$. By the continuity of $\mathcal{P}$, $\|\nabla J(\mathcal{P}(\hat{v}))\| \geq \delta > 0$. On the other hand, adding up two sides of (3.1) gives

$$\lim_{k \to \infty} J(u_k) - J(u_0) \leq -\frac{1}{4}\theta\alpha \sum_{k=0}^{\infty} s_k \|\nabla J(\mathcal{P}(v_k))\| \leq -\frac{1}{4}\theta\alpha\delta \sum_{k=0}^{\infty} s_k,$$

or $s_k \to 0$ as $k \to \infty$. It contradicts assumption (H). Therefore, there is a subsequence $\{v_{k_i}\}$ such that $\|\nabla J(\mathcal{P}(v_{k_i}))\| \to 0$ as $i \to \infty$ and $\{J(\mathcal{P}(v_{k_i}))\}$ is convergent. By the PS condition, $\{\mathcal{P}(v_{k_i})\}$ has a subsequence that converges to a critical point u_0.

(b) Suppose $u_0 = \mathcal{P}(v_0)$ is not a critical point. Then there is $\delta > 0$ such that $\|\nabla J(u_{k_i})\| > \delta, i = 1, 2, ...$ Similar to (3.1), we have

$$J(u_{k_i+1}) - J(u_{k_i}) \leq -\frac{1}{4}\theta\alpha s_{k_i} \|\nabla J(u_{k_i})\| < -\frac{1}{4}\theta\alpha\delta s_{k_i}.$$

Since $\sum_{k=0}^{\infty}[J(u_{k+1}) - J(u_k)] = \lim_{k \to \infty} J(u_k) - J(u_0)$, it leads to $\lim_{i \to \infty}(J(u_{k_i+1}) - J(u_{k_i})) = 0$. Hence $\lim_{i \to \infty} s_{k_i} = 0$. It contradicts assumption (H). Thus u_0 is a critical point. ∎

Next we prove an abstract existence-convergence result, that is actually independent of the algorithm and also explains why function values always converge faster than the gradients do. Denote $K_c = \{u \in B | \nabla J(u) = 0, J(u) = c\}$. By the PS condition, K_c is compact.

THEOREM 3.2. *Let $B = L \oplus L'$, $V \subset B$ be open and $U = V \cap S_{L'} \neq \emptyset$. Assume that $J \in C^1(B, \mathbb{R})$ satisfies the PS condition,*

(1) *$\mathcal{P}$ is a continuous L-$\perp$ selection of J in $\bar{U}$, where $\bar{U}$ is the closure of U on $S_{L'}$,*

(2) *$\inf_{v \in U} d(\mathcal{P}(v), L) > \alpha > 0$,*

(3) *$\inf_{v \in \partial\bar{U}} J(\mathcal{P}(v)) > c = \inf_{v \in U} J(\mathcal{P}(v)) > -\infty$, where $\partial\bar{U}$ is the boundary of $\bar{U}$ on $S_{L'}$.*

Then $K_c^p = \mathcal{P}(U) \cap K_c \neq \emptyset$ and for any $\{v_k\} \subset U$ with $J(u_k) \to c$ where $u_k = \mathcal{P}(v_k)$,

(a) *$\forall \varepsilon > 0$, there is $\bar{k} > 0$ such that $d(K_c^p, u_k) < \varepsilon$, $\forall k > \bar{k}$;*

(b) *If in addition, $\nabla J(\mathcal{P}(\cdot))$ is Lipschitz continuous in U, then there is a constant C such that $\|\nabla J(u_k)\| \leq C(J(u_k) - c)^{\frac{1}{2}}$.*

Proof. Define

$$\hat{J}(\mathcal{P}(v)) = \begin{cases} J(\mathcal{P}(v)) & v \in \bar{U}, \\ +\infty & v \notin \bar{U}. \end{cases}$$

Then, $\hat{J}(\mathcal{P}(\cdot))$ is lower semicontinuous and bounded from below on the complete metric space $S_{L'}$. Let $\{v_k\} \subset U$ be any sequence such that $J(\mathcal{P}(v_k)) \to c$. By our assumption (3), such a sequence always exists. Denote $u_k = \mathcal{P}(v_k)$. Applying Ekeland's variational principle to $\hat{J}(\mathcal{P}(\cdot))$, for every $v_k \in U$ and $\delta_k = (J(u_k) - c)^{\frac{1}{2}}$, there is $\bar{v}_k \in S_{L'}$ such that

$$(3.4) \qquad \hat{J}(\mathcal{P}(\bar{v}_k)) - \hat{J}(\mathcal{P}(v)) \leq \delta_k \|\bar{v}_k - v\|, \quad \forall v \in S_{L'},$$

$$(3.5) \qquad \hat{J}(\mathcal{P}(\bar{v}_k)) - \hat{J}(\mathcal{P}(v_k)) \leq -\delta_k \|\bar{v}_k - v_k\|.$$

By the definition of $\hat{J}(\mathcal{P}(\cdot))$ and assumptions on $\mathcal{P}$, we have $\bar{v}_k \in \bar{U}$,

$$(3.6) \qquad J(\mathcal{P}(\bar{v}_k)) - J(\mathcal{P}(v)) \leq \delta_k \|\bar{v}_k - v\|, \quad \forall v \in U,$$

$$(3.7) \qquad J(\mathcal{P}(\bar{v}_k)) - J(\mathcal{P}(v_k)) \leq -\delta_k \|\bar{v}_k - v_k\|.$$

It follows $c \leq J(\mathcal{P}(\bar{v}_k)) \leq J(u_k) - \delta_k \|\bar{v}_k - v_k\|$, or

$$(3.8) \qquad \|\bar{v}_k - v_k\| \leq \delta_k$$

and $d(L, \mathcal{P}(\bar{v}_k)) > \alpha$ when k is large. Then $J(\mathcal{P}(v_k)) \to c$ implies $J(\mathcal{P}(\bar{v}_k)) \to c$. By assumption (3), we have $\bar{v}_k \in U$ for large k. For those large k, if $\nabla J(\mathcal{P}(\bar{v}_k)) \neq 0$, by Lemma 2.4 and then Lemma 2.3, when s is small,

$$J(\mathcal{P}(\bar{v}_k(s))) - J(\mathcal{P}(\bar{v}_k)) \leq -\frac{\theta s}{4}|t_0^k| \|\nabla J(\mathcal{P}(\bar{v}_k))\| \leq -\frac{\alpha\theta}{8M}\|\nabla J(\mathcal{P}(\bar{v}_k))\| \|\bar{v}_k(s) - \bar{v}_k\|$$

where $\bar{v}_k(s) = \frac{\bar{v}_k + s\bar{w}_k}{\|\bar{v}_k + s\bar{w}_k\|} \in U$, $\bar{w}_k = -sign(t_0^k)G(\mathcal{P}(\bar{v}_k)))$, $\mathcal{P}(\bar{v}_k) = t_0^k\bar{v}_k + u_L^k$ for some $u_L^k \in L$, $G(\mathcal{P}(\bar{v}_k))$ is an L'-PPG of J at $\mathcal{P}(\bar{v}_k)$ with $\|G(\mathcal{P}(\bar{v}_k))\| \leq M$, see Lemma 2.2 and $|t_0^k| > \alpha$ by our assumption (2). Hence by (3.6), we get

$$(3.9) \qquad \|\nabla J(\mathcal{P}(\bar{v}_k))\| \leq \frac{8M}{\alpha\theta}\delta_k,$$

which implies $\nabla J(\mathcal{P}(\bar{v}_k)) \to 0$ and then $\nabla J(\mathcal{P}(v_k)) \to 0$ by (3.8). $\{J(\mathcal{P}(v_k))\}$ is already bounded. By the PS condition, $\{u_k\}$ has a subsequence that converges to a critical point u^*. By assumptions (3) and (1), it is clear that $u^* \in K_c^p \neq \emptyset$. Let β be any limit point of

$\{d(K_c^p, u_k)\}$ and $u_{k_i} = \mathcal{P}(v_{k_i}) \in \{u_k\}$ such that $\lim_{i \to \infty} d(K_c^p, u_{k_i}) = \beta$. By the PS condition, $\{\mathcal{P}(v_{k_i})\}$ has a subsequence that converges to a critical point $\bar{u}$. Again $\bar{u} \in K_c^p$, i.e., $\beta = 0$. Thus conclusion (a) holds.

If in addition, $\nabla J(\mathcal{P}(\cdot))$ is Lipschitz continuous in U with a Lipschitz constant ℓ_1, then by (3.8) and (3.9), conclusion (b) follows from

$$
\begin{aligned}
\|\nabla J(\mathcal{P}(v_k))\| &\leq \|\nabla J(\mathcal{P}(\bar{v}_k))\| + \|\nabla J(\mathcal{P}(v_k)) - \nabla J(\mathcal{P}(\bar{v}_k))\| \\
&\leq \frac{16M}{\alpha\theta}\delta_k + \ell_1\|\bar{v}_k - v_k\| \leq (\frac{16M}{\alpha\theta} + \ell_1)(J(u_k) - c)^{\frac{1}{2}}. \quad \blacksquare
\end{aligned}
$$

COROLLARY 3.1. *Let $J \in C^1(B, \mathbb{R})$ satisfy the PS condition, V_1 and V_2 be open in L' with $\emptyset \neq U_2 \equiv V_2 \cap S_{L'} \subset V_1 \cap S_{L'} \equiv U_1$. If $\mathcal{P}$ is a continuous L-$\perp$ selection of J in U_1 with*

(1) *$\inf_{v \in U_1} d(\mathcal{P}(v), L) \geq \alpha > 0$, $c = \inf_{v \in U_1} J(\mathcal{P}(v)) > -\infty$ and $K_c^p = \mathcal{P}(U_1) \cap K \subset K_c$,*

(2) *there is $d > 0$ with $\inf\{J(\mathcal{P}(v))|v \in U_1, d(v, \partial U_1) \leq d\} = a > b = \sup\{J(\mathcal{P}(v))|v \in U_2\}$,*

(3) *given $\{v_k\}$ such that $v_1 \in U_2$, $\|v_{k+1} - v_k\| < d$, $J(u_{k+1}) < J(u_k)$ and $\{u_k\}$ has a subsequence that converges to a critical point u_0, where $u_k = \mathcal{P}(v_k)$. Then*

(a) *$\forall \varepsilon > 0$, there is $\bar{k} > 0$ such that $d(K_c^p, u_k) < \varepsilon$, $\forall k > \bar{k}$;*

(b) *If in addition, $\nabla J(\mathcal{P}(\cdot))$ is Lipschitz continuous in U_1, then there is a constant C such that $\|\nabla J(u_k)\| \leq C(J(u_k) - c)^{\frac{1}{2}}$.*

Proof. First, we prove that $v_k \in U_1$ and $d(v_k, \partial U_1) > d$, $k = 1, 2, ...$ In fact, if $v_k \in U_1$, $d(v_k, \partial U_1) > d$ and $J(u_k) \leq b$, then $v_{k+1} \in U_1$ and $J(u_{k+1}) < b$, i.e., $v_{k+1} \in U_1$ and $d(v_{k+1}, \partial U_1) > d$. Thus, for $v_1 \in U_2$, $v_k \in U_1$ and $d(v_k, \partial U_1) > d$, $k = 1, 2, ...$ Since $K_c^p = \mathcal{P}(U_1) \cap K \subset K_c$ and $\{u_k\}$ has a subsequence that converges to a critical point u_0, we have $u_0 \in K_c^p \neq \emptyset$. Denote $U = \{v \in U_1 | d(v, \partial U_1) > d\}$. Then by the monotonicity of $\{J(u_k)\}$, we have $J(u_k) \to c = \inf_{v \in U} J(\mathcal{P}(v))$ as $k \to \infty$, and

$$
\inf_{v \in \partial \bar{U}} J(\mathcal{P}(v)) \geq a > b \geq J(\mathcal{P}(v_1)) \geq c = \inf_{v \in U} J(\mathcal{P}(v)).
$$

Thus all the assumptions of Theorem 3.2 are satisfied and the conclusions follow. $\blacksquare$

REMARK 3.1. There are two types of conditions posed in our convergence results. One is used in the literature to guarantee the existence of multiple solutions. The other is what we posed to insure a convergence for the algorithm. We will focus on verification of the later.

(1) In Theorem 3.2, if $U = S_{L'}$, condition (3) can be simplified as $c = \inf\limits_{v \in S_{L'}} J(\mathcal{P}(v)) > -\infty$.

(2) In Corollary 3.1, condition (2) or its variants are frequently used in the literature to form a topological linking for applying a deformation lemma to prove an existence of multiple solutions. It is clear that condition (3) in Theorem 3.2 is much weaker. It is used to trap a descending flow away from critical points at other levels. Condition (3) in Corollary 3.1 is designed for our algorithm to cover several different cases in Banach spaces and is guaranteed by our assumption (H) and Theorem 3.1.

(3) Note that when K_c^p contains only one point, Theorem 3.2 can be easily stated as a point-to-point convergence result. Theorem 3.2 together with its Corollary 3.1 improve a convergence result in Hilbert spaces, Theorem 3.3 in [5] in several directions. They (a) cover several different cases in Banach spaces, (b) do not require $\mathcal{P}$ be a homeomorphism and (c) contain a new result on relative convergence rate, i.e., inequality (2) which explains why in our numerical computations, $J(\mathcal{P}(v_n^k))$ always converges much faster than $\|\nabla J(\mathcal{P}(v_n^k))\| \to 0$.

Next we verify assumption (H) for several different cases. It is done in Lemmas 3.3, 3.4, 3.6 and 3.7 below. Cases 1 and 2 are general, so we assume $B = L \oplus L'$ where L' is a closed subspace of B and $\mathbb{P} : B \to L'$ is the corresponding projection. In Step 3 of the algorithm, we choose $w_k = -sign(t_0)G(\mathcal{P}(v_k))$ where G is either an L'-PPG of J or the value of an L'-PPG flow of J. Then $\|w_k\| \leq M = \|\mathbb{P}\|$. By Lemma 2.4 we obtain

LEMMA 3.2. *If $\mathcal{P}$ is a local L-$\perp$ selection of J at $v \in S_{L'}$ such that (1) $\mathcal{P}$ is continuous at v, (2) $d(\mathcal{P}(v), L) > 0$ and (3) $\nabla J(\mathcal{P}(v)) \neq 0$, then $s(v) > 0$.*

Case 1: Use the value of a negative PPG flow G as a descent direction.

Here $G(\mathcal{P}(v_k))$ is the value of an L'-PPG flow of J at $\mathcal{P}(v_k) = t_0^k v_k + v_k^L$ for some $v_k^L \in L$.

LEMMA 3.3. *If $\mathcal{P}$ is a local L-$\perp$ selection of J at $v_0 \in S_{L'}$ such that (1) $\mathcal{P}$ is continuous at v_0, (2) $d(\mathcal{P}(v_0), L) > 0$ and (3) $\nabla J(\mathcal{P}(v_0)) \neq 0$, then there exist $\varepsilon > 0$ and $s_0 = \frac{\lambda}{2^m}$ for*

some integer m such that for each $v \in S_{L'}$ with $\|v - v_0\| < \varepsilon$, $\lambda \geq s_0\|G(\mathcal{P}(v))\|$ and

$$J(\mathcal{P}(v(s_0))) - J(\mathcal{P}(v)) < -\frac{s_0\theta|t_v|}{4}\|\nabla J(\mathcal{P}(v))\|, \quad v(s_0) = \frac{v + sign(t_v)s_0 G(\mathcal{P}(v))}{\|v + sign(t_v)s_0 G(\mathcal{P}(v))\|},$$

$\mathcal{P}(v) = t_v v + w_v$ for some $w_v \in L$ and $G(\mathcal{P}(v))$ is the value of an L'-PPG flow of J at $\mathcal{P}(v)$

w.r.t. the constant θ.

Proof. By Lemma 2.4, there is $\bar{s} > 0$ such that as $0 < s < \bar{s}$

$$(3.10) \qquad J(\mathcal{P}(v_0(s))) - J(\mathcal{P}(v_0)) < -\frac{s\theta|t_0|}{4}\|\nabla J(\mathcal{P}(v_0))\|$$

where

$$v_0(s) = \frac{v_0 - sign(t_0)sG(\mathcal{P}(v_0))}{\|v_0 - sign(t_0)sG(\mathcal{P}(v_0))\|}, \quad \mathcal{P}(v_0) = t_0 v_0 + w_0$$

for some $w_0 \in L$. Actually, for each fixed s, the two sides of (3.10) are continuous in v_0.

Thus, there are $\varepsilon > 0$, $s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda \geq s_0\|G(\mathcal{P}(v))\|$ and

$$J(\mathcal{P}(v(s_0))) - J(\mathcal{P}(v)) < -\frac{s_0\theta|t_v|}{4}\|\nabla J(\mathcal{P}(v))\|,$$

$\forall v \in S_{L'}$ with $\|v - v_0\| \leq \varepsilon$. ∎

Case 2: Use a negative PPG G as a descent direction.

Here $G(\mathcal{P}(v_k))$ is a L'-PPG of J at $\mathcal{P}(v_k) = t_0^k v_k + v_k^L$ for some $v_k^L \in L$. Since an L'-PPG may be chosen from different L'-PPG flows, we lost the continuity. To compensate the loss, we assume that an L-$\perp$ selection $\mathcal{P}$ of J is locally Lipschitz continuous.

LEMMA 3.4. *Let $\mathcal{P}$ be a local L-$\perp$ selection of J at $v_0 \in S_{L'}$. If (1) $\mathcal{P}$ is Lipschitz continuous in a neighborhood of v_0, (2) $d(\mathcal{P}(v_0), L) > 0$ and (3) $\nabla J(\mathcal{P}(v_0)) \neq 0$, then there are $\varepsilon > 0$ and $s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda \geq s_0\|G(\mathcal{P}(v))\|$ and*

$$J(\mathcal{P}(v(s_0))) - J(\mathcal{P}(v)) < -\frac{1}{4}s_0\theta|t_v|\|\nabla J(\mathcal{P}(v))\|,$$

$\forall v \in S_{L'}$ with $\|v - v_0\| < \varepsilon$, where

$$v(s) = \frac{v - sign(t_v)sG(\mathcal{P}(v))}{\|v - sign(t_v)sG(\mathcal{P}(v))\|}, \quad s > 0, \quad \mathcal{P}(v) = t_v v + w_v \text{ for some } w_v \in L$$

and $G(\mathcal{P}(v))$ is an L'-PPG of J at $\mathcal{P}(v)$ w.r.t. the constant θ.

Proof. First, denote $\mathcal{P}(v(s)) = t_v^s v(s) + w_v(s)$ for some $w_v(s) \in L$, we have

$$(3.11)\, J(\mathcal{P}(v(s))) - J(\mathcal{P}(v)) = \langle \nabla J(\mathcal{P}(v)) + (\nabla J(\zeta(v,s)) - \nabla J(\mathcal{P}(v))), \mathcal{P}(v(s)) - \mathcal{P}(v) \rangle$$

where $\zeta(v,s) = (1-\eta)\mathcal{P}(v) + \eta\mathcal{P}(v(s))$ for some $\eta \in [0,1]$. By assumption (1) and Lemma 2.3, as s is small and for any v close to v_0,

$$(3.12) \qquad \|\mathcal{P}(v(s)) - \mathcal{P}(v)\| \le \ell\|v(s) - v\| \le \frac{2\ell s\|G(\mathcal{P}(v))\|}{\|v - sign(t_v)sG(\mathcal{P}(v))\|} \le 4\ell M s.$$

On the other hand, by the definition of an L-$\perp$ selection of J, we have

$$
\begin{aligned}
\langle \nabla J(\mathcal{P}(v)), \mathcal{P}(v(s)) - \mathcal{P}(v)\rangle &= -\frac{sign(t_v)t_v^s s\langle \nabla J(\mathcal{P}(v)), G(\mathcal{P}(v))\rangle}{\|v - sign(t_v)sG(\mathcal{P}(v))\|} \\
(3.13) \qquad = -\frac{|t_v^s|s\langle \nabla J(\mathcal{P}(v)), \Psi(\mathcal{P}(v))\rangle}{\|v - sign(t_v)sG(\mathcal{P}(v))\|} &\le -\frac{s\theta|t_v|\|\nabla J(\mathcal{P}(v))\|}{2} < 0
\end{aligned}
$$

and

$$
\begin{aligned}
&|\langle \nabla J(\zeta(v,s)) - \nabla J(\mathcal{P}(v)), \mathcal{P}(v(s)) - \mathcal{P}(v)\rangle| \\
(3.14) \qquad &\le \|\nabla J(\zeta(v,s)) - \nabla J(\mathcal{P}(v))\| \, \|\mathcal{P}(v(s)) - \mathcal{P}(v)\| \le \frac{s\theta|t_v|\|\nabla J(\mathcal{P}(v))\|}{4}
\end{aligned}
$$

where in the last inequality, since J is C^1 and by assumptions (2) and (3), we have

$$(3.15) \qquad \|\nabla J(\zeta(v,s)) - \nabla J(\mathcal{P}(v))\| \le \frac{\theta|t_v|\|\nabla J(\mathcal{P}(v))\|}{16\ell M}.$$

By (3.11) and the boundedness of L'-PPGs, there exist $\varepsilon > 0$ and $s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda \ge s_0\|G(\mathcal{P}(v))\|$ and

$$J(\mathcal{P}(v(s_0))) - J(\mathcal{P}(v)) \le -\frac{s_0\theta|t_v|\|\nabla J(\mathcal{P}(v))\|}{4}, \ \forall v \in S_{L'} \text{ with } \|v - v_0\| < \varepsilon. \quad \blacksquare$$

Case 3: Use a practical technique for a descent direction in $B = W_0^{1,p}(\Omega)$.

Let $B = W_0^{1,p}(\Omega)$ where $\Omega \subset \mathbb{R}^n$ is open and bounded, $p > 1$, $B^* = W^{-1,q}(\Omega)$ with $\frac{1}{p} + \frac{1}{q} = 1$. The usual gradient $\delta J(u) \in B^* = W^{-1,q}(\Omega)$ cannot be used as a search direction. Thus $d = \Delta_p^{-1}(\delta J(u)) \in B$ has been used in the literature as a descent direction to find a local minimum of $J : B \to \mathbb{R}$, where Δ_p is the p-Laplacian operator defined in (4.1) and Δ_p^{-1} is its inverse. It leads to solve a sequence of quasi-linear elliptic equations $\Delta_p d_k = \delta J(u_k)$. But such d is not a PPG, it does not help much for finding a saddle point. A practical technique is used in [12] for numerical implementation to compute a PPG. The results are very promising. Here we wish to provide some mathematical justification. This technique is based on the understanding that when a nice smooth initial guess v_0 is used,

we may expect that 'nice' functions are actually used to approximate a critical point. Let $\mathcal{P}$ be an L-$\perp$ selection of J. For $v \in S_{L'}, u = \mathcal{P}(v)$, by the definition of $\mathcal{P}, \delta J(u) \perp L$. But $\delta J(u) \in W^{-1,q}(\Omega)$, its smoothness is poor. We first lift its smoothness by computing $d := \nabla J(u) = \Delta^{-1}(-\delta J(u)) \in W_0^{1,q}(\Omega)$, i.e., $d_k = \nabla J(u_k)$ is solved from

$$\Delta d_k(x) = -\delta J(u_k)(x), \ x \in \Omega, \quad d_k(x)|_{\partial\Omega} = 0.$$

Observe that notationally for any $w \in B$,

$$\langle d, w \rangle_{W_0^{1,q} \times W_0^{1,p}} \equiv \langle \nabla d, \nabla w \rangle_{L^q \times L^p} \equiv \int_\Omega \nabla d(x) \cdot \nabla w(x)\, dx$$

$$= \int_\Omega -\Delta d(x) w(x)\, dx = \int_\Omega \delta J(u)(x) w(x)\, dx \equiv \langle \delta J(u), w \rangle_{W^{-1,q} \times W_0^{1,p}}.$$

This suggests that $d = \nabla J(u)$ be used as a gradient of J at u. In particular when $u = \mathcal{P}(v)$,

(3.16) $$\langle \nabla J(u), w \rangle_{W_0^{1,q} \times W_0^{1,p}} = \langle \delta J(u), w \rangle_{W^{-1,q} \times W_0^{1,p}} = 0, \quad \forall w \in [L, v],$$

or $\nabla J(u) \perp [L, v]$. This suggests a natural way to choose L'. We will discuss it late. Since

$$\|\delta J(u)\|_{W^{-1,q}} = \sup_{\|w\|_{W_0^{1,p}}=1} \langle \delta J(u), w \rangle_{W^{-1,q} \times W_0^{1,p}}$$

(3.17) $$= \sup_{\|w\|_{W_0^{1,p}}=1} \langle d, w \rangle_{W_0^{1,q} \times W_0^{1,p}} = \sup_{\|\nabla w\|_{L^p}=1} |\langle \nabla d, \nabla w \rangle_{L^q \times L^p}| \leq \|d\|_{W_0^{1,q}},$$

the PS condition of J in terms of δJ implies the PS condition of J in terms of ∇J. In our convergence analysis of the algorithm, the first order approximation contains a term

$$\langle \delta J(v_0), \nabla J(v_0) \rangle_{W^{-1,q} \times W_0^{1,p}} = \int_\Omega \delta J(v_0) \nabla J(v_0) dx$$

$$= \int_\Omega -\Delta(\nabla J(v_0)) \nabla J(v_0) dx = \int_\Omega \nabla(\nabla J(v_0)) \cdot \nabla(\nabla J(v_0)) dx = \|\nabla J(v_0)\|_{W^{1,2}}^2,$$

which will be used to design a new stepsize rule. Next we let $u_k = \mathcal{P}(v_k)$ and check the ratio

(3.18) $$1 \geq \theta_k \equiv \frac{\|\nabla J(u_k)\|_2^2}{\|\nabla J(u_k)\|_q \|\nabla J(u_k)\|_p} \geq \theta > 0 \quad \forall k = 1, 2, ...,$$

where $\| \cdot \|_r$ is the norm in $W_0^{1,r}(\Omega)$ with $r > 1$. Let $G(u_k) = \dfrac{\nabla J(u_k)}{\|\nabla J(u_k)\|_p}$. We have $\|G(u_k)\|_p \leq M = 1$ and

(3.19) $$\langle \delta J(u_k), G(u_k) \rangle = \frac{\|\nabla J(u_k)\|_2^2}{\|\nabla J(u_k)\|_p} = \theta_k \|\nabla J(u_k)\|_q \geq \theta_k \|\delta J(u_k)\|_{W^{-1,q}} \geq \theta \|\delta J(u_k)\|_{W^{-1,q}}$$

where the last inequality holds if (3.18) is satisfied, i.e., $G(u_k)$ is a pseudo-gradient of J at u_k. Then (3.16) suggests that $G(u_k)$ is also an L'-PPG of J at $u_k = \mathcal{P}(v_k)$ if $L' = L^\perp$ where $L^\perp$ is given in (3.20), $L = [w_1, ..., w_{n-1}]$ and $w_1, ..., w_{n-1}$ are linearly independent. To show $B = L \oplus L^\perp$, we need further assume that when $1 < p < 2$, $w_1, ..., w_{n-1}$ are n-1 previously found 'nice' critical points, or at least they are 'nice' approximations of some exact critical points such that $L \subset W_0^{1,q}(\Omega)$. Such an assumption holds automatically when $2 \le p$. Thus

$$(3.20) \qquad L' := L^\perp = \{u \in B | \int_\Omega \nabla u(x) \cdot \nabla v(x) \, dx = 0, \ \forall v \in L\}$$

is well defined and $L \cap L' = \{0\}$ holds. For any $w \in B$, we compute $w_L := \sum_{i=1}^{n-1} \alpha_i w_i$ from

$$\int_\Omega \nabla w_L(x) \cdot \nabla w_j(x) \, dx = \int_\Omega \nabla w(x) \cdot \nabla w_j(x) \, dx, \ \ j = 1, ..., n-1.$$

Thus $w_L \in L$ and $w - w_L \in L^\perp$, i.e., $B = L \oplus L'$.

But we cannot assume that such $G(u_k)$ is the value of a PPG flow of J at $u_k = \mathcal{P}(v_k)$, because we do not know the ratio at other points. In all our numerical examples, (3.18) is satisfied. But we note that the ratio is stable for $1 < p \le 2$ and gets closer to 0 as $p \to +\infty$. Thus we treat those two cases differently in our convergence analysis. For $1 < p \le 2$, we assume that (3.18) is satisfied. But for $p > 2$, we only assume $\|\nabla J(u_k)\|_p \le M$ for some $M > 0$. Either one of the assumptions implies $\nabla J(u_k) \in L^\perp \subset B$. By comparing $G(u_k)$ and $\nabla J(u_k)$, Step 3 and the stepsize rule in Step 5 need to be modified as below.

Step 3: Find a descent direction w_k of J at $u_k = \mathcal{P}(v_k)$, $w_k = -sign(t_0^k)\nabla J(u_k)$. Compute the ratio $\theta_k = \dfrac{\|w_k\|_2^2}{\|w_k\|_p \|w_k\|_q} > 0$;

Since $\langle \delta J(u_k), \nabla J(u_k) \rangle = \|\nabla J(u_k)\|_2^2$, the stepsize rule in Step 5 has to be changed to

$$(3.21) \quad s_k = \max\left\{s = \frac{\lambda}{2^m} | m \in N, 2^m > \|w_k\|, J(\mathcal{P}(v_k(s))) - J(u_k) \le \frac{|t_0^k|s}{-4}\|\nabla J(u_k)\|_2^2\right\},$$

where $0 < \lambda < 1$. Note that if $\theta_k > \theta > 0$, theoretically the term $\|\nabla J(u_k)\|_2^2$ in the above stepsize rule can be replaced by $\theta\|\nabla J(u_k)\|_q$, i.e., we use $G(\mathcal{P}(v_k)) = \dfrac{\nabla J(u_k)}{\|\nabla J(u_k)\|_p}$ as an L'-PPG of J at $u_k = \mathcal{P}(v_k)$. Then this case can be covered by Case 2. But in implementation, the lower bound θ of the ratio is usually not known beforehand. In particular, we do not known whether or not the ratio is satisfied at a limit point of the sequence. Thus the current stepsize rule (3.21) has to be used in implementation. First we show that if $0 < \|\nabla J(\mathcal{P}(v_0))\|_p < +\infty$, then a positive stepsize can always be attained.

LEMMA 3.5. *For $v_0 \in S_{L'}$, if J has a local L-$\perp$ selection $\mathcal{P}$ at v_0 satisfying (1) $\mathcal{P}$ is continuous at v_0, (2) $d(\mathcal{P}(v_0), L) > 0$ and (3) $0 < \|\nabla J(\mathcal{P}(v_0))\|_2 < +\infty$. Then there exists $s_0 > 0$ such that as $0 < s < s_0$*

$$(3.22) \qquad J(\mathcal{P}(v_0(s))) - J(\mathcal{P}(v_0)) < -\frac{|t_0|s}{4}\|\nabla J(\mathcal{P}(v_0))\|_2^2$$

where $\mathcal{P}(v_0) = t_0 v_0 + w_0$ for some $t_0 \in \mathbb{R}$, $w_0 \in L$ and $v_0(s) = \dfrac{v_0 - sign(t_0)s\nabla J(\mathcal{P}(v_0))}{\|v_0 - sign(t_0)s\nabla J(\mathcal{P}(v_0))\|}$.

Proof. Since $\|\mathcal{P}(v_0(s)) - \mathcal{P}(v_0)\| \to 0$ as $s \to 0$, we have

$$J(\mathcal{P}(v_0(s))) - J(\mathcal{P}(v_0)) = \langle \delta J(\mathcal{P}(v_0)), \mathcal{P}(v_0(s)) - \mathcal{P}(v_0)\rangle + o(\|\mathcal{P}(v_0(s)) - \mathcal{P}(v_0)\|)$$

$$= -\frac{sign(t_0)t_0^s s\|\nabla J(\mathcal{P}(v_0))\|_2^2}{\|v_0 - sign(t_0)s\nabla J(\mathcal{P}(v_0))\|} + o(\|\mathcal{P}(v_0(s)) - \mathcal{P}(v_0)\|) < -\frac{|t_0|s}{4}\|\nabla J(\mathcal{P}(v_0))\|_2^2$$

where $\mathcal{P}(v_0(s)) = t_0^s v_0(s) + w_0^s$ and $w_0^s \in L$, when $s > 0$ is very small. ∎

Next we verify assumption (H).

Subcase p<2. (We assume (3.18) holds.) We have $\nabla J(u_k) \in W_0^{1,q}(\Omega) \subset B$. The conclusion in the next lemma is actually stronger than assumption (H).

LEMMA 3.6. *Let $J \in C^1(B, \mathbb{R})$ and $v_0 \in S_{L^\perp}$. Let $\mathcal{P}$ be a local L-$\perp$ selection of J at v_0 such that $\mathcal{P}$ is continuous at v_0 and $d(\mathcal{P}(v_0), L) > 0$. If $\delta J(\mathcal{P}(v_0)) \neq 0$, then there are $\varepsilon > 0$ and $s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda \geq s_0\|\nabla J(\mathcal{P}(v))\|_p$ and*

$$J(\mathcal{P}(v(s_0))) - J(\mathcal{P}(v)) < -\frac{|t_v|s_0}{4}\|\nabla J(\mathcal{P}(v))\|_2^2,$$

$\forall v \in S_{L'}, \|v - v_0\| < \varepsilon$, where $\mathcal{P}(v) = t_v v + w$, $w \in L$ and $v(s_0) = \dfrac{v - sign(t_v)s_0\nabla J(\mathcal{P}(v))}{\|v - sign(t_v)s_0\nabla J(\mathcal{P}(v))\|}$.

Proof. By Lemma 3.5, we have

$$(3.23) \qquad J(\mathcal{P}(v_0(s))) - J(\mathcal{P}(v_0)) < -\frac{|t_0|s}{4}\|\nabla J(\mathcal{P}(v_0))\|_2^2.$$

When $p < 2$, we have $q > 2$. J is C^1 implies that ∇J is continuous in $\|\cdot\|_2$-norm. For fixed s, all the terms in (3.23) are continuous in v_0. Thus there exists $\varepsilon > 0$ and $s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda \geq s_0\|\nabla J(\mathcal{P}(v))\|_p$ since J is C^1 and

$$J(\mathcal{P}(v(s_0))) - J(\mathcal{P}(v)) < -\frac{|t_v|s_0}{4}\|\nabla J(\mathcal{P}(v))\|_2^2, \ \forall v \in S_{L'}, \|v - v_0\| < \varepsilon. \quad ∎$$

With the new stepsize rule and the uniform stepsize result, Lemma 3.6, if $\theta_k > \theta > 0$ holds in Step 3. We can verify Theorem 3.1. The proof is similar. We need only replace (3.1) by

$$J(u_{k+1}) - J(u_k) \; < \; -\frac{\alpha s_k}{4}\|\nabla J(u_k)\|_2^2 \quad \text{(by (3.18))}$$

$$(3.24) \qquad < \; -\frac{\alpha\theta s_k}{4}\|\nabla J(u_k)\|_p\|\nabla J(u_k)\|_q < -\frac{\alpha\theta|1-\lambda|}{4M}\|v_{k+1}-v_k\|\|\nabla J(u_k)\|_q$$

where $0 < \lambda < 1$ is given in (3.21) and then follow the proof.

Subcase p>2. (We only assume $\|\nabla J(u_k)\|_p \leq M$ for some $M > 0$.) We have $B = W_0^{1,p}(\Omega) \subset W_0^{1,2}(\Omega)$. To verify assumption (H) and prove the convergence of the algorithm, we note that in this case, $\nabla J(u_k) \in L' = L^\perp$ still holds, i.e., $-\nabla J(u_k)$ can be used as a search direction. But J is C^1 means that δJ is continuous in $\|\cdot\|_{(-1,q)}$-norm and ∇J is continuous in $\|\cdot\|_q$-norm, not necessarily in $\|\cdot\|_2$-norm. Thus we need an L-$\perp$ selection $\mathcal{P}$ to be locally Lipschitz continuous.

LEMMA 3.7. *Let $J \in C^1(B,\mathbb{R})$ and $v_0 \in S_{L'}$. Assume that $\mathcal{P}$ is a local L-$\perp$ selection of J at v_0 such that (1) $\mathcal{P}$ is locally Lipschitz continuous (2) $d(\mathcal{P}(v_0),L) > 0$ and (3) $\delta J(\mathcal{P}(v_0)) \neq 0$. Then for any $v_k \in S_{L'}$ with $\lim_{k\to\infty} v_k = v_0$ and $\|\nabla J(\mathcal{P}(v_k))\|_p \leq M$, there are $\bar{k}, s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda > s_0\|\nabla J(\mathcal{P}(v_k))\|_p$ and*

$$J(\mathcal{P}(v_k(s_0))) - J(\mathcal{P}(v_k)) \leq -\frac{s_0|t_k|\|\nabla J(\mathcal{P}(v_k))\|_2^2}{4}, \quad \forall k > \bar{k},$$

where $v_k(s) = \dfrac{v_k - sign(t_k)s\nabla J(\mathcal{P}(v_k))}{\|v_k - sign(t_k)s\nabla J(\mathcal{P}(v_k))\|}$ and $\mathcal{P}(v_k) = t_k v_k + v_k^L$, $v_k^L \in L$.

Proof. Denote $\mathcal{P}(v_k(s)) = t_k^s v_k(s) + v_k^L(s)$ for some $v_k^L(s) \in L$. Then, by the mean value theorem, we have

$$(3.25) \quad J(\mathcal{P}(v_k(s))) - J(\mathcal{P}(v_k)) = \langle \delta J(\mathcal{P}(v_k)) + (\delta J(\zeta(v_k,s)) - \delta J(\mathcal{P}(v_k))), \mathcal{P}(v_k(s)) - \mathcal{P}(v_k)\rangle$$

where $\zeta(v_k,s) = (1-\lambda_k)\mathcal{P}(v_k) + \lambda_k\mathcal{P}(v_k(s))$ for some $\lambda_k \in [0,1]$. By assumption (1) and Lemma 2.3,

$$\|\mathcal{P}(v_k(s)) - \mathcal{P}(v_k)\| \leq \ell\|v_k(s) - v_k\| \leq \frac{2\ell s\|\nabla J(\mathcal{P}(v_k))\|_p}{\|v_k - sign(t_k)s\nabla J(\mathcal{P}(v_k))\|}.$$

On the other hand, by the definition of an L-$\perp$ selection of J, as $s > 0$ is small and k is large, we have

$$\langle \delta J(\mathcal{P}(v_k)), \mathcal{P}(v_k(s)) - \mathcal{P}(v_k)\rangle \; = \; -\frac{sign(t_k)t_k^s s\|\nabla J(\mathcal{P}(v_k))\|_2^2}{\|v_k - sign(t_k)s\nabla J(\mathcal{P}(v_k))\|}$$

$$(3.26) \qquad\qquad\qquad\qquad \leq \; -\frac{s|t_k|}{2}\|\nabla J(\mathcal{P}(v_k))\|_2^2 < 0.$$

Since J is C^1 and $1 < q < 2$ in this case, by assumptions (2), (3) and applying inequality (3.17), there exists $\delta > 0$ such that when s is small and k is large,

$$\frac{|t_k| \|v_k - sign(t_k) s \nabla J(\mathcal{P}(v_k))\| \|\nabla J(\mathcal{P}(v_k))\|_2^2}{8\ell \|\nabla J(\mathcal{P}(v_k))\|_p} > \delta > 0.$$

Thus we can choose $s > 0$ small and k large such that

$$\|\delta J(\zeta(v_k, s)) - \delta J(\mathcal{P}(v_k))\| \leq \frac{|t_k| \|v_k - sign(t_k) s \nabla J(\mathcal{P}(v_k))\| \|\nabla J(\mathcal{P}(v_k))\|_2^2}{8\ell \|\nabla J(\mathcal{P}(v_k))\|_p}.$$

Hence

$$|\langle \delta J(\zeta(v_k, s)) - \delta J(\mathcal{P}(v_k)), \mathcal{P}(v_k(s)) - \mathcal{P}(v_k) \rangle|$$

(3.27) $\quad \leq \quad \|\delta J(\zeta(v_k, s)) - \delta J(\mathcal{P}(v_k))\| \|\mathcal{P}(v_k(s)) - \mathcal{P}(v_k)\| \leq \dfrac{s|t_k| \|\nabla J(\mathcal{P}(v_k))\|_2^2}{4}.$

Applying inequalities (3.26) and (3.27) to (3.25), there exist $\bar{k}, s_0 = \frac{\lambda}{2^m}$ for some integer m such that $\lambda > s_0 \|\nabla J(\mathcal{P}(v_k))\|_p$ and

$$J(\mathcal{P}(v_k(s_0))) - J(\mathcal{P}(v_k)) \leq -\frac{s_0 |t_k| \|\nabla J(\mathcal{P}(v_k))\|_2^2}{4}, \ \forall k > \bar{k}. \quad \blacksquare$$

With the new stepsize rule (3.21) and the assumption $\|\nabla J(u_k)\|_p < M$, the conclusion of Lemma 3.7 implies assumption (H), i.e., $s(v_k) \geq s_0$. Then the convergence result, Theorem 3.1, can be verified. The proof is similar. Note that when $\|\nabla J(u_k)\|_q > \delta_0$ for some $\delta_0 > 0$, $\|\nabla J(u_k)\|_2 > \delta$ for some $\delta > 0$. We only need to replace (3.1) and (3.2) by

$$\begin{aligned} J(u_{k+1}) - J(u_k) \quad &< \quad -\frac{\alpha s_k}{4} \|\nabla J(u_k)\|_2^2 \ \leq \ -\frac{\alpha s_k}{4} \delta^2 \\ &\leq \quad -\frac{\alpha s_k \delta^2}{4M} \|\nabla J(u_k)\|_p \ \leq \ -\frac{\alpha \delta^2 |1 - \lambda|}{4M} \|v_{k+1} - v_k\|_p \end{aligned}$$

where $0 < \lambda < 1$ is given in (3.21) and the last inequality follows from Lemma 2.3. Then following the proof, the unified convergence result, Theorem 3.1 is also obtained.

4   An Application to Nonlinear p-Laplacian Equation

As an application, let us consider the following quasilinear elliptic boundary-value problem on a bounded smooth domain $\Omega \subset \mathbb{R}^n$

(4.1) $\quad \begin{cases} \Delta_p u(x) + f(x, u(x)) \ = \ 0, \quad x \in \Omega, \\ \qquad\qquad\quad u(x) \ = \ 0, \quad x \in \partial\Omega, \end{cases} \quad u \in B \equiv W^{1,p}(\Omega), \ p > 1,$

where Δ_p defined by $\Delta_p u(x) = div(|\nabla u(x)|^{p-2}\nabla u(x))$ is the p-Laplacian operator which has a variety of applications in physical fields, such as in fluid dynamics when the shear stress and the velocity gradient are related in certain manner where $p = 2, p < 2, p > 2$ if the fluid is Newtonian, pseudoplastic, dilatant, respectively. The p-Laplacian operator also appears in the study of flow in a porous media $(p = \frac{3}{2})$, nonlinear elasticity $(p > 2)$ and glaciology $(p \in (1, \frac{4}{3}))$. Under certain standard conditions on f, it can be shown that a point $u^* \in W_0^{1,p}(\Omega)$ is a weak solution of (4.1) if and only if u^* is a critical point of the functional

$$(4.2) \quad J(u) = \frac{1}{p}\int_\Omega |\nabla u(x)|^p dx - \int_\Omega F(x, u(x))dx \quad \text{where} \quad F(x,t) = \int_0^t f(x,s)ds.$$

Many multiple solutions to the above quasilinear elliptic PDE have been numerically computed in [12,13] for $p < 2$ and $p > 2$. Convergence results obtained in Section 3 can be applied, see Case 3. Since conditions (1), (2) and (3) in Theorem 3.1 are basic assumptions in our results and new in the literature, we focus on verifying them in this section. While other conditions, such as the PS condition, have been studied in the literature and therefore will not be discussed here. Let us assume some of the standard growth and regularity conditions in the literature. Denote the Sobolev exponent $p^* = \frac{np}{n-p}$ for $p < n$ and $p^* = \infty$ for $p \geq n$.

Assume

(a) $f \in C^1(\bar{\Omega} \times \mathbb{R}, \mathbb{R})$, $f(x, 0) = 0$, $\dfrac{f(x, t\xi)}{|t\xi|^{p-2}t\xi}$ monotonically increases to $+\infty$ in t,

(b) For each $\varepsilon > 0$, there is $c_1 = c_1(\varepsilon) > 0$ such that $f(x, t)t < \varepsilon|t|^p + c_1|t|^{p^*}$ $\forall t \in \mathbb{R}$, $x \in \Omega$.

It is clear that $u = 0$ is a critical point of the least critical value of J and $f(x, u) = |u|^{q-2}u$ for $q > p$ satisfies condition (a). For each $v \in B$ with $\|v\| = 1$ and $t > 0$, let $g(t) = J(tv)$. We have

$$
\begin{aligned}
g'(t) &= \langle \nabla J(tv), v\rangle = \int_\Omega \left(t^{p-1}|\nabla v(x)|^p - f(x, tv(x))v(x)\right)dx \\
&= t^{p-1}\left(1 - \int_\Omega \frac{f(x, tv(x))|v(x)|^p}{|tv(x)|^{p-2}tv(x)}\right)dx.
\end{aligned}
$$

Thus, by condition (a), there is a unique $t_v > 0$ such that $g'(t_v) = 0$, i.e., for $L = \{0\}$ and each $v \in S_B$, the L-$\perp$ selection (actually a peak selection) $\mathcal{P}(v) = t_v v$ is uniquely determined with $J(\mathcal{P}(v)) > 0$. By taking a derivative of condition (a) w.r.t. t, we have

$$
\begin{aligned}
g''(t) &= (p-1)t^{(p-2)} - \int_\Omega f'_u(x, tv(x))v^2(x)\, dx \\
(4.3) \qquad &< (p-1)t^{(p-2)} - \int_\Omega \frac{(p-1)}{t}f(x, tv(x))v(x)\, dx = \frac{p-1}{t}g'(t).
\end{aligned}
$$

Thus condition (3) in Theorem 3.1 is always satisfied for any L. Next let us recall that when $L = [u^1, u^2, ..., u^{n-1}]$, by the definition of an L-$\perp$ selection, $\mathcal{P}(v) = t_0 v + t_1 u^1 + \cdots + t_{n-1} u^{n-1}$ is solved from

$$(4.4) \quad \frac{\partial}{\partial t_0} g(t_0, ..., t_{n-1}) = \langle \nabla J(t_0 v + t_1 u^1 + \cdots + t_{n-1} u^{n-1}), v \rangle = 0,$$

$$\frac{\partial}{\partial t_i} g(t_0, ..., t_{n-1}) = \langle \nabla J(t_0 v + t_1 u^1 + \cdots + t_{n-1} u^{n-1}), u^i \rangle = 0, \quad i = 1, ..., n-1,$$

where $g(t_0, ..., t_{n-1}) = J(t_0 v + + t_1 u^1 + ... + t_{n-1} u^{n-1})$. If $u = \mathcal{P}(v) = t_0 v + t_1 u^1 + \cdots + t_{n-1} u^{n-1}$ satisfies (4.4) and at u, the $n \times n$ matrix

$$Q = \left[\frac{\partial^2}{\partial t_i \partial t_j} g(t_0, ..., t_{n-1}) \right] = \begin{bmatrix} \langle J''(u)v, v \rangle & \langle J''(u)u^1, v \rangle & \cdots & \langle J''(u)u^{n-1}, v \rangle \\ \langle J''(u)v, u^1 \rangle & \langle J''(u)u^1, u^1 \rangle & \cdots & \langle J''(u)u^{n-1}, u^{n-1} \rangle \\ \cdots & & \cdots & \\ \langle J''(u)v, u^{n-1} \rangle & \langle J''(u)u^1, u^{n-1} \rangle & \cdots & \langle J''(u)u^{n-1}, u^{n-1} \rangle \end{bmatrix}$$

is invertible, i.e., $|Q| \neq 0$, then by the implicit function theorem, around u, the L-$\perp$ selection $\mathcal{P}$ is well-defined and continuously differentiable. The condition $|Q| \neq 0$ can be easily and numerically checked. For the current case $L = \{0\}$, by (4.3), we have $Q = g''(t_v) < 0$. Thus the L-$\perp$ selection $\mathcal{P}$ is C^1. To show that $d(\mathcal{P}(v), L) > \alpha > 0$ for all $v \in S_B$, by (b), for any $\varepsilon > 0$, there is $c_1 = c_1(\varepsilon)$ such that $f(x, v(x))v(x) < \varepsilon |v(x)|^p + c_1 |v(x)|^{p^*}$. It follows

$$\int_\Omega f(x, v(x))v(x) \, dx < \varepsilon \int_\Omega |v(x)|^p \, dx + c_1 \int_\Omega |v(x)|^{p^*} \, dx$$

(by the Poincare and Sobolev inequalities)

$$\leq \varepsilon c_0(\Omega) \int_\Omega |\nabla v(x)|^p \, dx + c_1 c_2(\Omega) \left(\int_\Omega |\nabla v(x)|^p \, dx \right)^{\frac{p^*}{p}}$$

$$= \left[\varepsilon c_0(\Omega) + c_1 c_2(\Omega) \left(\int_\Omega |\nabla v(x)|^p \, dx \right)^{\frac{p^*}{p} - 1} \right] \int_\Omega |\nabla v(x)|^p \, dx.$$

Thus

$$\langle \nabla J(v), v \rangle \geq \left[1 - \varepsilon c_0(\Omega) - c_1 c_2(\Omega) \left(\int_\Omega |\nabla v(x)|^p \, dx \right)^{\frac{p^*}{p} - 1} \right] \int_\Omega |\nabla v(x)|^p \, dx$$

$$= \left[1 - \varepsilon c_0(\Omega) - c_1 c_2(\Omega) \|v\|^{p^* - p} \right] \|v\|^p.$$

It follows that for any small $\varepsilon > 0, c_1, c_0(\Omega)$ and $c_2(\Omega)$, there is $t_0 > 0$ such that when $0 < \|v\| = t < t_0$, we have $\langle \nabla J(v), v \rangle \geq \left[1 - \varepsilon c_0(\Omega) - c_1 c_2(\Omega) t^{p^* - p} \right] t^p > 0$. Therefore the L-$\perp$ selection $\mathcal{P}(v)$ satisfies $\|\mathcal{P}(v)\| > t_0$ or $d(\mathcal{P}(v), L) > t_0 > 0, \forall v \in S_B$ where $L = \{0\}$.

References

[1] Y.S.CHOI AND P.J.MCKENNA, *A mountain pass method for the numerical solution of semilinear elliptic problems*, Nonlinear Analysis, Theory, Methods and Applications, 20(1993), 417-437.

[2] Z.DING, D.COSTA AND G.CHEN, *A high linking method for sign changing solutions for semilinear elliptic equations*, Nonlinear Analysis, 38(1999), 151-172.

[3] PREM K. KYTHE, *Differential Operators and applications* , Birkhauser, Boston, 1996.

[4] Y.LI AND J.ZHOU, *A minimax method for finding multiple critical points and its applications to nonlinear PDEs*, SIAM Sci. Comp., 23(2001), 840-865.

[5] Y.LI AND J.ZHOU, *Convergence results of a local minimax method for finding multiple critical points*, SIAM Sci. Comp., 24(2002), 840-865.

[6] F.LIN AND T.LIN, *Minimax solutions of the Ginzburg-Landau equations*, Slecta Math.(N.S.), 3(1977) no.1, 99-123.

[7] J.MAWHIN AND M.WILLEM, *Critical Point Theory and Hamiltonian Systems*, Springer-Verlag, New York, 1989.

[8] P.RABINOWITZ, *Minimax Method in Critical Point Theory with Applications to Differential Equations*, CBMS Regional Conf. Series in Math., No.65, AMS, Providence, 1986.

[9] MARTIN SCHECHTER, *Linking Methods in Critical Point Theory*, Birkhauser, Boston, 1999.

[10] M.STRUWE, *Variational Methods*, Springer, 1996.

[11] M.WILLEM, *Minimax Theorems*, Birkhauser, Boston, 1996.

[12] X.YAO AND J.ZHOU, *A Minimax Method for Finding Multiple Critical Points in Banach Spaces and Its Application to Quasilinear Elliptic PDE*, SIAM J. Sci. Comp., 26(2005), 1796-1809.

[13] X.YAO AND J.ZHOU, *Numerical Methods for Computing Nonlinear Eigenpairs: Part I. Iso-Homogeneous Cases*, revised and in review.

[14] J.ZHOU, *A Local Min-Orthogonal Method for Finding Multiple Saddle Points*, J. Math. Anal. Appl., 291(2004), 66-81.