A Local Minimax Characterization for Computing Multiple Nonsmooth Saddle Critical Points *

Xudong Yao[†] and Jianxin Zhou[†]

July 29, 2004

Abstract

This paper is concerned with characterizations of nonsmooth saddle critical points for numerical algorithm design. Most characterizations for nonsmooth saddle critical points in the literature focus on existence issue and are converted to solve global minimax problems. Thus they are not helpful for numerical algorithm design. Inspired by the results on computational theory and methods for finding multiple smooth saddle critical points in [14, 15, 19, 21, 23], a local minimax characterization for multiple nonsmooth saddle critical points in either a Hilbert space or a reflexive Banach space is established in this paper to provide a mathematical justification for numerical algorithm design. A local minimax algorithm for computing multiple nonsmooth saddle critical points is presented by its flow chart.

Keywords. Locally Lipschitz Continuous Functional, Nonsmooth Critical Points, Minimax Characterization

AMS(MOS) subject classifications. 58E05, 58E30,35A40,35J65

Abbreviated titles. Nonsmooth Minimax Characterization

1 Introduction

Let B be a Banach space, B^* its dual space, $\langle , \rangle$ the dual relation and $\|\cdot\|$ its norm. Let $J : B \to R$ be a locally Lipschitz continuous functional. Then the generalized gradient $\partial J(u)$

*This research is supported in part by NSF Grant DMS-0311905

[†]Math. Dept., Texas A&M University, College Station, TX 77843 (xdyao@math.tamu.edu, jzhou@math.tamu.edu)

of J at $u \in B$ in the sense of Clark [5] is defined as follows.

DEFINITION 1.1. *Let J be Lipschitz continuous near $u_0 \in B$. The generalized directional derivative $J^0(u_0; v)$ of J at u_0 in the direction of $v \in B$ is defined by*

$$J^0(u_0; v) = \limsup_{\substack{u \to u_0 \\ t \downarrow 0}} \frac{J(u_0 + tv) - J(u_0)}{t}.$$

The generalized gradient $\partial J(u_0)$ of J at u_0 is a subset of B^* given by

$$\partial J(u_0) = \{\zeta \in B^* : J^0(u_0; v) \geq \langle \zeta, v \rangle, \forall v \in B\}.$$

Then according to Chang [4], a point $u^* \in B$ is a critical point of J iff

$$0 \in \partial J(u^*). \tag{1.1}$$

If J is C^1, (1.1) reduces to $\nabla J(u^*) = 0$ where $\nabla J(u^*)$ is the gradient of J at u^*, i.e., (1.1) becomes the wellknown Euler-Lagrange equation. The first candidates for critical points are the local maxima or minima to which the classical optimization theory, calculus of variation and nonsmooth analysis were devoted. Traditional numerical methods focus on finding such stable solutions. Critical points u^* that are not local extrema are called *saddle points*, i.e., u^* satisfies (1.1) and for any neighborhood $\mathcal{N}(u^*)$ of u^* there are $v, w \in \mathcal{N}(u^*)$ such that

$$J(v) < J(u^*) < J(w).$$

In physical systems, saddle points appear as unstable equilibria or transient excited states.

Critical points of a C^1 functional are called smooth critical points and critical points of a locally Lipschitz continuous functional are called nonsmooth critical points.

For smooth critical points, the Mountain Pass Lemma established in 1973 by Ambrosetti-Rabinowitz [1] set a milestone in contemporary critical point theory. Since then the study of smooth critical points has gained a tremendous popularity and various saddle point and linking theorems were established in the literature to prove existence of multiple critical points for various nonlinear problems [1,3,4,12,13,17]. However all those theorems focus on existence issue and are based on a deformation lemma or a global minimax approach, i.e., the problem is converted to find a solution to

$$\min_{A\in\mathcal{A}}\max_{u\in A} J(u)$$

where $\mathcal{A}$ is usually a class of compact sets in B and both min and max are global. Therefore they are not helpful to devise numerical algorithms for finding multiple critical points. So a new kind of approach in critical point theory must be developed. A local minimax approach in Hilbert spaces was established by Li-Zhou in [14,15]. Based on a local minimax characterization of multiple saddle points, a local minimax algorithm was developed and successfully implemented to solve semilinear elliptic equations for multiple solutions. With the introduction of a modified pseudo-gradient, Yao-Zhou [19] extended the results in [14, 15] to Banach spaces and successfully solved many quasilinear elliptic equations for multiple solutions [19] and nonlinear eigenpair problems [21]. Some convergence results of the local minimax methods are also established by Li-Zhou and Yao-Zhou in [15, 20].

The notion of nonsmooth critical points was first introduced by Chang [4] in 1981, who obtained a nonsmooth version of the saddle point theorem of Rabinowitz by proving a nonsmooth version of deformation lemma. Kourogenis and Papageorgious [12] generalized Chang's results. Kourogenis-Kandilakis-Papageorgious [13] obtained a nonsmooth version of Linking Theorem. All those results and others [2,11, 18] are used to prove existence of multiple nonsmooth saddle points for problems in differential inclusion, hemivariational inequalities and eigenvalue problems with discontinuities. Similar to the smooth critical point cases, those results in the literature focus on the existence issue and are based on a nonsmooth version of deformation lemma or a global minimax approach, and therefore not helpful to devise numerical algorithms.

The key to our success in finding multiple smooth critical points is the establishment of a local minimax characterization of multiple saddle points. This inspires us in this paper to try to establish a local minimax characterization of multiple nonsmooth saddle points.

In the rest of this section, we introduce some notions we need for this paper.

In Section 2.1, we establish a local minimax characterization for multiple nonsmooth saddle critical points in Hilbert spaces which generalize the results in [14, 15]. The purpose of Section 2.2 is to establish a local minimax characterization for multiple nonsmooth saddle critical points in Banach spaces. To do so, first we have to introduce the definition of a pseudo-generalized-gradient, a new notion in critical point theory. As an application, a local minimax algorithm for finding multiple nonsmooth critical points will be presented by its flow chart in Section 3. An example is constructed and studied in Section 4 to illustrate the new theory and method. We are conducting further study and implementation of the algorithm.

A typical application of nonsmooth critical point theory to partial differential equations is the Dirichlet problem:

$$\begin{cases} -\Delta_p u(x) = f(u(x)), & \text{in } \Omega, \\ u(x) = 0, & \text{on } \partial\Omega. \end{cases} \tag{1.2}$$

where Ω is an open bounded domain in R^n with smooth boundary $\partial\Omega$ and $\Delta_p u(x) = \text{div}(|\nabla u(x)|^{p-2}\nabla u(x))$ is the nonlinear p-Laplacian differential operator and f is a locally bounded measurable function. The corresponding variational functional

$$J(u) = \frac{1}{p}\int_\Omega |\nabla u(x)|^p dx - \int_\Omega F(u(x))dx$$

where $F(t) = \int_0^t f(s)ds$, is locally Lipschitz continuous and then critical points of J correspond to weak solutions of (1.2).

1.1 Some Definitions and Preliminaries

For a subspace $B' \subseteq B$, denote $S_{B'} = \{v : v \in B', \|v\| = 1\}$ the unit sphere in B'. Assume that $B = L \oplus L'$ where L and L' are two closed subspaces of B. For each $v \in S_{L'}$, denote $[L, v]$ the subspace spanned by v and L, and denote $\mathcal{P}$ the corresponding projection operator from B to L'.

DEFINITION 1.2. *A set-valued mapping $P : S_{L'} \to 2^B$ is the peak mapping of J if $\forall v \in S_{L'}, P(v) = \{u \in [L, v] :$ u is a local maximum point of J in $[L, v]\}$. A single-valued mapping $p : S_{L'} \to B$ is a peak selection of J if $p(v) \in P(v)$, $\forall v \in S_{L'}$. For a given $v \in S_{L'}$, if p is locally defined in a neighborhood of a point $v \in S_{L'}$, we say that J has a local peak selection p at v.*

REMARK 1.1. (a) The above definition of a peak selection is an important notion in our local minimax approach for multiple saddle points. The purpose of introducing a peak

selection is to provide a separation from previously found saddle critical points that span L.
(b) In a Banach space B, if L is finite dimensional, then there always exists a closed subspace $L' \subset B$ such that $B = L \oplus L'$, [8].
(c) If B is a Hilbert space, then L' is usually chosen to be $L^\perp$.

Let us recall some basic lemmas for the generalized gradient of locally Lipschitz continuous functionals which will be used later for convenience.

Lemma 1.1. *([5]) Assume that J is Lipschitz continuous in a neighborhood $\mathcal{N}(u_0)$ of u_0 with Lipschitz constant K, i.e., $|J(u) - J(v)| \le K\|u - v\|$, $\forall u, v \in \mathcal{N}(u_0)$. Then*

(1) *For all $u \in \mathcal{N}(u_0)$, $\partial J(u)$ is a nonempty, convex, weak*-compact subset of B^* and $\|w\| \le K$, $\forall w \in \partial J(u)$.*

(2) *Let B be a Hilbert space. For each $u \in \mathcal{N}(u_0)$, if $z \in \partial J(u)$ such that $\|z\| = \min\{\|\zeta\| : \zeta \in \partial J(u)\}$, then we have*

$$\langle z, \zeta\rangle \ge \|z\|^2, \quad \forall \zeta \in \partial J(u).$$

Lemma 1.2. *(Lebourg, [5]) Let $u, v \in B$. Assume that J is Lipschitz continuous in an open set which contains the line segment $\{\lambda u + (1-\lambda)v : \lambda \in [0,1]\}$. Then there is $\lambda_0 \in (0,1)$ such that*

$$J(u) - J(v) \in \langle \partial J(\lambda_0 u + (1-\lambda_0)v), u - v\rangle.$$

Definition 1.3. *A set-valued mapping $G : B \to 2^{B^*}$ is said to be weakly upper semicontinuous at $u \in B$, if $u_k \to u$ and $v_k \in G(u_k)$ there is $w_k \in G(u)$ such that $w_k - v_k \to 0$ weakly. G is said to be weakly upper semicontinuous if it is weakly upper semicontinuous at each point in B. If $w_k - v_k \to 0$ strongly then G is said to be upper semicontinuous.*

2 Local Minimax Characterization for Multiple Nonsmooth Critical Points

2.1 Local Minimax Characterization in Hilbert Spaces

For easy understanding, let us first consider the case in a Hilbert space H. By using the generalized gradient, we are able to establish a local minimax characterization for multiple nonsmooth critical points in H which generalizes the corresponding results in [14, 15] for multiple smooth critical points in H. The following lemma plays an important role in the local minimax method.

LEMMA 2.1. *Let H be a Hilbert space with $H = L \oplus L^{\perp}$ for a closed subspace $L \subset H$ and $J : H \to \mathbb{R}$. Assume that p is a local peak selection of J w.r.t. L at $v \in S_{L^{\perp}}$ and J is locally Lipschitz continuous in a neighborhood of $p(v)$ such that*

(1) *p is continuous at v and $dis(p(v), L) > 0$,*

(2) *$z \in \partial J(p(v))$ such that $\|z\| = \min\{\|w\| : w \in \partial J(p(v))\} > 0$.*

Then as $s > 0$ is sufficient small,

$$J(p(v(s))) - J(p(v)) < -\frac{1}{4}|t_v|\|z\|^2, \tag{2.1}$$

where $v(s) = \dfrac{v - sign(t_v)sz_{L^{\perp}}}{\|v - sign(t_v)sz_{L^{\perp}}\|}$, $p(v) = t_v v + w_v$, $w_v \in L$ and $z = z_L + z_{L^{\perp}}$, $z_L \in L$, $z_{L^{\perp}} \in L^{\perp}$.

Proof. By Lemma 1.2, for t close to t_v, $w \in L$ close to w_v and $s > 0$ sufficient small,

$$\begin{aligned} & J\Big(\frac{tv}{\|v - sign(t_v)sz_{L^{\perp}}\|} + w - \frac{sign(t_v)stz}{\|v - sign(t_v)sz_{L^{\perp}}\|}\Big) - J\Big(\frac{tv}{\|v - sign(t_v)sz_{L^{\perp}}\|} + w\Big) \\ = & -\frac{sign(t_v)st}{\|v - sign(t_v)sz_{L^{\perp}}\|}\langle z_{v,w}, z\rangle \end{aligned}$$

where $z_{v,w} \in \partial J\Big(\dfrac{tv}{\|v - sign(t_v)sz_{L^{\perp}}\|} + w - \lambda_{v,w}\dfrac{sign(t_v)stz}{\|v - sign(t_v)sz_{L^{\perp}}\|}\Big)$ for some $\lambda_{v,w} \in (0,1)$. Since p is a peak selection, for t close to t_v and w close to w_v, we have

$$J(p(v)) \geq J\Big(\frac{tv}{\|v - sign(t_v)sz_{L^{\perp}}\|} + w\Big).$$

Hence

$$J\Big(\frac{tv}{\|v - sign(t_v)sz_{L^{\perp}}\|} + w - \frac{sign(t_v)stz}{\|v - sign(t_v)sz_{L^{\perp}}\|}\Big) - J(p(v)) \leq -\frac{sign(t_v)st}{\|v - sign(t_v)sz_{L^{\perp}}\|}\langle z_{v,w}, z\rangle.$$

On the other hand, the set-valued mapping $u \to \partial J(u)$, is weakly upper semicontinuous [6][1], there is $\zeta_{v,w} \in \partial J(p(v))$ such that

$$|\langle \zeta_{v,w} - z_{v,w}, z\rangle| \leq \frac{1}{2}\|z\|^2$$

[1]The authors would like to thank an anonymous reviewer for pointing out this result and the reference.

for t close to t_v, w close to w_v and $s>0$ sufficient small. Thus, by Lemma 1.1,

$$\begin{aligned} & J(\frac{tv}{\|v-sign(t_v)sz_{L^\perp}\|}+w-\frac{sign(t_v)stz}{\|v-sign(t_v)sz_{L^\perp}\|})-J(p(v)) \\ \leq \; & -\frac{sign(t_v)st}{\|v-sz_{L^\perp}\|}(-|\langle z_{v,w}-\zeta_{v,w},z\rangle|+\langle \zeta_{v,w},z\rangle) \leq -\frac{1}{4}s|t_v|\|z\|^2. \end{aligned}$$

Then

$$J(p(v(s)))-J(p(v))<-\frac{1}{4}s|t_v|\|z\|^2,$$

as $s>0$ sufficient small by letting $t=t(s)$, $w=w(s)+\dfrac{sign(t_v)st(s)z_L}{\|v-sign(t_v)sz_{L^\perp}\|}$, where $p(v(s))=t(s)v(s)+w(s)$, $w(s)\in L$. ∎

Remark 2.1. (a) $z_{L^\perp}\neq 0$ since $z\neq 0$ and p is a peak selection, see Lemma 3.1.
(b) The inequality (2.1) is an important result which can be used to not only derive a local minimax characterization of nonsmooth saddle point as presented in Theorem 2.1 but also design a stepsize rule for the local minimax algorithm, see Step 5 in the flow chart of the algorithm in Section 3. ∎

By Lemma 2.1, a minimax characterization for nonsmooth critical points in a Hilbert space can be immediately established as follow.

Theorem 2.1. *Let H be a Hilbert space with $H=L\oplus L^\perp$ for a closed subspace $L\subset H$ and $J:H\to\mathbb{R}$. Assume that p is a local peak selection of J w.r.t. L at $v\in S_{L^\perp}$ and J is locally Lipschitz continuous in a neighborhood of $p(v)$ such that*

(1) *p is continuous at v and $dis(p(v),L)>0$,*

(2) *$J(p(v))=local\text{-}min_{u\in S_{L^\perp}}J(p(u))$.*

Then $p(v)$ is a critical point of J.

Proof. If $p(v)$ is not a critical point of J, let $z\in\partial J(p(v))$ satisfying $\|z\|=\min\{\|w\|:w\in\partial J(p(v))\}>0$, then by Lemma 2.1, as $s>0$ sufficient small,

$$J(p(v(s)))-J(p(v))<-\frac{1}{4}s|t_v|\|z\|^2,$$

where $v(s)=\dfrac{v-sign(t_v)sz_{L^\perp}}{\|v-sign(t_v)sz_{L^\perp}\|}$, $p(v)=t_vv+w_v$, $w_v\in L$ and $z=z_L+z_{L^\perp}$, $z_L\in L$, $z_{L^\perp}\in L^\perp$. It is a contradiction to assumption (2). ∎

2.2 Local Minimax Characterization in Reflexive Banach Spaces

Now we start to establish a local minimax characterization for nonsmooth saddle points in Banach spaces. Since in this case, the generalized gradient $\partial J(u)$ is in B^* not B, a point in $\partial J(u)$ cannot be used to update an iteration point $u \in B$. Thus as long as numerical algorithms are concerned, a new notion has to be developed. Motivated by the notion of a pseudo-gradient for smooth critical points in Banach spaces, we introduce the following definition which is crucial for later development.

DEFINITION 2.1. *Let B be a reflexive Banach space and $J : B \to \mathbb{R}$ be Lipschitz continuous near a point $u_0 \in B$. Let $\mu = \min\{\|z\|_{B^*} : z \in \partial J(u_0)\}$. Then the pseudo-generalized-gradient (PGG) $\Psi J(u_0)$ of J at u_0 is defined by*

$$\Psi J(u_0) = \{z^* \in B : \|z^*\| = \mu, \inf_{w\in\partial J(u_0)} \langle w, z^*\rangle \geq \langle z, z^*\rangle = \mu^2, z \in \partial J(u_0), \|z\|_{B^*} = \mu\}. \tag{2.2}$$

LEMMA 2.2. *Assume that B is a reflexive Banach space and $J : B \to \mathbb{R}$ is Lipschitz continuous near a point $u_0 \in B$. Then the PGG $\Psi J(u_0)$ of J at u_0 is a nonempty, convex set in B. If in addition, B^* is locally uniformly convex and $\|\cdot\|_{B^*}$ is Frechet differentiable on $B^* \setminus \{0\}$, then $\Psi J(u_0) = \{\|z\|_{B^*}\|z\|'_{B^*}\}$ where z is the unique point of minimum norm in $\partial J(u_0)$.*

Proof. Let $\mu = \min\{\|z\|_{B^*} : z \in \partial J(u_0)\}$ and $S(\mu) = \{u \in B^* : \|u\|_{B^*} \leq \mu\}$. If $0 \in \partial J(u_0)$, i.e., $\mu = 0$ and $z = 0$, then $\Psi J(u_0) = \{0\}$. If $0 \notin \partial J(u_0)$, then $\mu > 0$ and there is $z \in \partial J(u_0)$ such that $\|z\|_{B^*} = \mu > 0$ since $\partial J(u_0)$ is convex and weak*-compact. Note that $\text{int}S(\mu) \cap \partial J(u_0) = \emptyset$ and $z \in S(\mu) \cap \partial J(u_0)$, by Lemma 1.1 and the separation theorem [22], there is a $z^* \in B^{**} = B$ such that

(1) $\|z^*\| = \|z\|_{B^*} = \mu$, and

(2) $\inf_{w\in\partial J(u_0)}\langle w, z^*\rangle = \langle z, z^*\rangle = \sup_{u\in S(\mu)}\langle u, z^*\rangle$.

On the other hand,

$$\sup_{u\in S(\mu)} \langle u, z^*\rangle = \sup_{\{u\in B^*:\|u\|=\|z\|\}} \langle u, z^*\rangle = \|z^*\|\|z\|_{B^*} = \|z\|^2_{B^*} = \mu^2.$$

Hence

$$\langle w, z^*\rangle \geq \langle z, z^*\rangle = \|z\|^2_{B^*} = \mu^2, \quad \forall w \in \partial J(u_0).$$

Thus $\Psi J(u_0)$ is nonempty. To show that $\Psi J(u_0)$ is convex, let $z_1^*, z_2^* \in \Psi J(u_0)$ and $0 < \alpha < 1$. There exist $z_1, z_2 \in \partial J(u_0)$ such that $\|z_1\|_{B^*} = \|z_2\|_{B^*} = \mu > 0$ and

$$\langle w, z_i^* \rangle \geq \langle z_i, z_i^* \rangle = \mu^2, \quad \forall w \in \partial J(u_0),\ i = 1, 2.$$

Since $\|z_1^*\| = \|z_2^*\| = \mu$, we have

$$\|\alpha z_1^* + (1-\alpha) z_2^*\| \leq \mu \text{ and } \|\alpha z_1 + (1-\alpha) z_2\| \leq \mu,$$

and for all $w \in \partial J(u_0)$,

$$\begin{aligned} & \langle w, \alpha z_1^* + (1-\alpha) z_2^* \rangle = \alpha \langle w, z_1^* \rangle + (1-\alpha) \langle w, z_2^* \rangle \\ \geq\ & \alpha \langle z_1, z_1^* \rangle + (1-\alpha) \langle z_2, z_2^* \rangle = \alpha \mu^2 + (1-\alpha) \mu^2 = \mu^2. \end{aligned}$$

In particular for $w = \alpha z_1 + (1-\alpha) z_2 \in \partial J(u_0)$, we have

$$\mu^2 \leq \langle \alpha z_1 + (1-\alpha) z_2, \alpha z_1^* + (1-\alpha) z_2^* \rangle \leq \|\alpha z_1 + (1-\alpha) z_2\|_{B^*} \|\alpha z_1^* + (1-\alpha) z_2^*\| \leq \mu^2.$$

Therefore we must have

$$\begin{aligned} \langle \alpha z_1 + (1-\alpha) z_2, \alpha z_1^* + (1-\alpha) z_2^* \rangle &= \mu^2, \\ \|\alpha z_1^* + (1-\alpha) z_2^*\| = \|\alpha z_1 + (1-\alpha) z_2\|_{B^*} &= \mu \end{aligned}$$

and for all $w \in \partial J(u_0)$,

$$\langle w, \alpha z_1^* + (1-\alpha) z_2^* \rangle \geq \langle \alpha z_1 + (1-\alpha) z_2, \alpha z_1^* + (1-\alpha) z_2^* \rangle = \mu^2,$$

i.e., $\alpha z_1^* + (1-\alpha) z_2^* \in \Psi J(u_0)$ and thus $\Psi J(u_0)$ is a convex set.

If in addition, B^* is locally uniformly convex and $\|\cdot\|_{B^*}$ is Frechet differentiable on $B^* \setminus \{0\}$, then there is only one $z \in \partial J(u_0)$ with $\|z\|_{B^*} = \mu$ and $S(\mu) \cap \partial J(u_0) = \{z\}$. The set $\{u \in B^* : \langle \|z\|'_{B^*}, u - z \rangle = 0\}$ is the tangent plane of the sphere $S(\mu)$ at z. On the other hand, B is reflexive and $\|\cdot\|'_{B^*}$ exists on on $B^* \setminus \{0\}$ imply that B is locally uniformly convex. Since $\Psi J(u_0)$ is a convex set in B such that for any $z^* \in \Psi J(u_0)$, we have $\|z^*\| = \|z\|_{B^*} = \mu$, the set $\Psi J(u_0)$ can contain at most one point z^*. The hyperplane corresponding to z^* separates $\partial J(u_0)$ from $S(\mu)$ at z. Such a separating hyperplane must be a tangent plane of $S(\mu)$ at z. Since $g(v) = \|v\|_{B^*}$ is Frechet differentiable at z, such a tangent plane is unique. We have

$$\langle \|z\|'_{B^*}, w - z\rangle \geq 0 \geq \langle \|z\|'_{B^*}, u - z\rangle, \forall w \in \partial J(u_0), u \in S(\mu).$$

Since $\langle \|z\|'_{B^*}, z\rangle = \|z\|_{B^*} = \mu$, we have $\forall w \in \partial J(u_0), u \in S(\mu)$,

$$\langle \|z\|_{B^*}\|z\|'_{B^*}, w\rangle \geq \langle \|z\|_{B^*}\|z\|'_{B^*}, z\rangle = \mu^2 = \|z\|^2_{B^*} \geq \langle \|z\|_{B^*}\|z\|'_{B^*}, u\rangle,$$

which implies $\|\|z\|_{B^*}\|z\|'_{B^*}\| = \|z\|_{B^*} = \mu$ and then $\Psi J(u_0) = \{z^*\} = \{\|z\|_{B^*}\|z\|'_{B^*}\}$. ∎

Remark 2.2. (a) When B is a Hilbert space, $z^* = z$.

(b) When J is a C^1 functional, z^* is a pseudo-gradient of J at u_0 with

$$\|z^*\| = \|\nabla J(u_0)\| \quad \textit{and} \quad \langle z^*, \nabla J(u_0)\rangle \geq \|\nabla J(u_0)\|^2.$$

(c) By the Kadec-Troyanski theorem (pp. 603-605, [22]), in every reflexive Banach space B, an equivalent norm $\|\cdot\|_B$ can be introduced so that B and B^* are locally uniformly convex and therefore $\|\cdot\|_B$ and $\|\cdot\|_{B^*}$ are Frechet differentiabl on $B \setminus \{0\}$ and $B^* \setminus \{0\}$. Thus in this case, we may use the norm $\|\cdot\|_B$ as the default norm $\|\cdot\|$ on B.

Then replacing the generalized gradient by the PGG and with some modification, the following lemma can be verified in a similar way as in Lemma 2.1.

Lemma 2.3. *Let B be a reflexive Banach space with $B = L \oplus L'$. Assume that J is locally Lipschitz continuous in B and p is a local peak selection of J w.r.t L at $v \in S_{L'}$ such that*

(1) *p is continuous at v and $dis(p(v), L) > 0$,*

(2) *$z^* \in B$ is the PGG of J at $p(v)$ with $\|z^*\| > 0$.*

Then

$$J(p(v(s))) - J(p(v)) < -\frac{1}{4}s|t_v|\|z\|^2_{B^*},$$

where $v(s) = \dfrac{v - sign(t_v)sz^*_{L'}}{\|v - sign(t_v)sz^*_{L'}\|}$, $p(v) = t_v v + w_v$, $w_v \in L$, $z^* = z^*_L + z^*_{L'}$, $z^*_L \in L$, $z^*_{L'} \in L'$ *and z is a point of minimum norm in $\partial J(p(v))$.*

By Lemma 2.3, the minimax characterization for nonsmooth critical points in Banach spaces can be written as follow.

THEOREM 2.2. *Let B be a reflexive Banach space with $B = L \oplus L'$. Assume that J is locally Lipschitz continuous in B and p is a local peak selection of J w.r.t L at $v \in S_{L'}$ such that*

(1) *p is continuous at v and $dis(p(v), L) > 0$,*

(2) *$J(p(v)) = local\text{-}min_{u \in S_{L'}} J(p(u))$.*

Then $p(v)$ is a critical point of J, i.e., $0 \in \partial J(p(v))$.

3 A Local Minimax Algorithm

Before we present the algorithm, we need the following lemma to show that Step 3 in the algorithm can be carried out once a nonsmooth saddle critical point has not been reached.

LEMMA 3.1. *Let B be a reflexive Banach space with $B = L \oplus L'$ for some closed subspaces L, L' in B and $J : B \to \mathbb{R}$. Assume p is a local peak selection of J w.r.t L at $v_0 \in S_{L'}$ and J is locally Lipschitz continuous near $u_0 = p(v_0)$. If u_0 is not a critical point, then $\mathcal{P}(z^*) \neq 0, \forall z^* \in \Psi J(u_0)$.*

Proof. Since u_0 is not a critical point of J, we have $\mu = \min\{\|z\| : z \in \partial J(u_0)\} > 0$. If $\mathcal{P}(z^*) = 0$, then $z^* \in L$, $u_0 + tz^* \in [L, v_0]$. When $t > 0$ is sufficiently small, by Lemma 1.2, there exist $\lambda \in (0, 1)$, $\zeta_t \in \partial J(u_0 + \lambda t z^*)$ and $\zeta_0 \in \partial J(u_0)$ such that

$$\begin{aligned} J(u_0 + tz^*) - J(u_0) &= t\langle \zeta_t, z^* \rangle = t(\langle \zeta_t - \zeta_0, z^* \rangle + \langle \zeta_0, z^* \rangle) \\ &\geq t(-\frac{1}{2}\mu^2 + \mu^2) = \frac{t}{2}\mu^2 > 0, \end{aligned}$$

where the first inequality is due to the fact that $G : u \to \partial J(u)$ is weakly upper semicontinuous and $z^* \in \Psi J(u_0)$. It leads to a contradiction to the assumption that $u_0 = p(v_0)$ is a local maximum point of J in $[L, v_0]$. ∎

Now we are ready to present the algorithm.

Assume that u_1,...,u_{n-1} are $n-1$ previously found nonsmooth critical points of a locally Lipschitz continuous functional J in a reflexive Banach space B. Let $L = \{u_1, ..., u_{n-1}\}$, $B = L \oplus L'$ and $\mathcal{P}$ be the corresponding projection operator from B to L'. Given $\varepsilon, \lambda > 0$.

A flow chart of the algorithm reads:

Step 1: Let $v_n^1 \in S_{L'}$ be an increasing-decreasing direction at u_{n-1}.

Step 2: Set $k=1$ and solve for

$$\begin{aligned} u_n^k &= p(v_n^k) = t_0^k v_n^k + t_1^k u_1 + \cdots + t_{n-1}^k u_{n-1} \\ &= \arg\max\{J(t_0 v_n^k + t_1 u_1 + \cdots + t_{n-1}u_{n-1}) | t_i \in R, i=0,1,...,n-1\}. \end{aligned}$$

Step 3: Find a descent direction $w_n^k = -\mathrm{sign}(t_0^k)\mathcal{P}(z_n^k)$ at u_n^k, where $z_n^k \in \Psi J(u_n^k)$.

Step 4: If $\|u_n^k - u_n^{k-1}\| < \varepsilon$, then output u_n^k, stop. Otherwise, do Step 5.

Step 5: For each $s>0$, let $v_n^k(s) = \dfrac{v_n^k + s w_n^k}{\|v_n^k + s w_n^k\|}$ and use the initial point $(t_0^k, t_1^k, ..., t_{n-1}^k)$ to solve for

$$p(v_n^k(s)) = \arg\max\left\{J(t_0 v_n^k(s) + \sum_{i=1}^{n-1} t_i u_i) | t_i \in R, i=0,1,...,n-1\right\},$$

then set $u_n^{k+1} = p(v_n^{k+1}) = p(v_n^k(s_n^k)) \equiv t_0^{k+1} v_n^{k+1} + \sum_{i=1}^{n-1} t_i^{k+1} u_i$ where s_n^k satisfies

$$s_n^k = \max\{s = \frac{\lambda}{2^m} | m \in N, 2^m > \|w_n^k\|, J(p(v_n^k(s))) - J(p(v_n^k)) \le -\frac{1}{4}|t_0^k| s \|z_n^k\|^2\}.$$

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

REMARK 3.1. (a) By Lemmas 2.3 and 3.1, a positive step size in Step 5 of the algorithm can always be obtained if a critical point has not been reached. Therefore the algorithm is a strict descending method, i.e., $J(u_n^k) < J(u_n^{k-1})$, $\forall k = 1, 2, ...$
(b) When B is a Hilbert space, L' will be chosen as $L^\perp$ and z_n^k is the unique point of minimum norm in $\partial J(u_n^k)$.
(c) When J is a C^1 functional, this algorithm will reduce to the local minimax algorithm in [14, 15] if B is a Hilbert space and the local minimax algorithm in [19] if B is a reflexive Banach space except Step 3 where for smooth saddle critical points [14, 15,19],

$$\|\nabla J(u_n^k)\| \le \varepsilon \quad \textit{or} \quad \|G_n^k\| \le \varepsilon,$$

where G_n^k is a modified pseudo-gradient of J at u_n^k, is naturally used as a criterion to stop iteration in the algorithm. For nonsmooth saddle critical points, one may think to use

$$\|\mathcal{P}(z_n^k)\| \leq \varepsilon \tag{3.1}$$

as a criterion to stop iteration in the algorithm. But it is easy to construct a Lipschitz continuous functional J, e.g., $J(u) = |u|$ for $u \in \mathbb{R}$, such that u_0 is a nonsmooth critical point of J and $u_n^k \to u_0 \in B$ satisfies

$$\|\mathcal{P}(z_n^k)\| > \beta > 0, \forall z_n^k \in \Psi J(u_n^k).$$

Hence in general (3.1) cannot be used as a criterion to stop iteration in the algorithm. Instead we may use $\|u_n^k - u_n^{k-1}\| < \varepsilon$ or $J(u_n^k) - J(u_n^{k-1}) < \varepsilon$ since the algorithm is strictly descending, or $\|v_n^k - v_n^{k-1}\| < \varepsilon$ which is equivalent to $\|s_n^k \mathcal{P}(z_n^k)\| < \varepsilon$, as a criterion to stop the iteration of the algorithm. Those criteria are commonly used in numerical computation.
(d) Other definitions of generalized gradient may also be used to derive local minimax characterization of nonsmooth saddle critical points. We are conducting further study and implementation of the algorithm.

4 An Example

To illustrate the theory and method presented in the previous sections, let us consider the following simple example. Let $H = H_0^1(\Omega)$ where $\Omega \subset \mathbb{R}^n$ is an open, bounded domain. Many semilinear elliptic boundary-valued problems (inclusions) can be converted to find a nonsmooth critical point of the functional J on H defined by

$$J(u) = \int_\Omega \left\{ \frac{1}{2} |\nabla u(x)|^2 - F(u(x)) \right\} dx \tag{4.1}$$

where F is a locally Lipschitz continuous function, e.g., $F(s) = \int_0^s f(t)dt$ and f is a measurable function [4,10]. Then under some standard conditions, J is a locally Lipschitz continuous functional. Assume

(a) $F(s) \geq 0$, $F(s) = o(s^2)$ as $s \to 0$ and $s \to \infty$,

(b) $\dfrac{\partial F(s)}{s}$ is monotone in the sense that for any $|s_2| > |s_1| > 0$, any two selections $\delta_1 F$ and $\delta_2 F$ of ∂F, we have

$$\frac{\delta_2 F(s_2)}{s_2} > \frac{\delta_1 F(s_1)}{s_1}.$$

REMARK 4.1. A smooth version of Condition (b) above has been frequently used in the literature [7]. Condition (a) is to ensure the possession of a mountain pass structure for a functional and Condition (b) is to guarantee the uniqueness of the peak along each direction.

THEOREM 4.1. *For $L = \{0\}, L' = H$, assume that Conditions (a) and (b) above are satisfied and F is regular [5] at every point. Then the peak mapping $P : S_{L'} \to 2^H$ is well-defined and single-value, i.e., for each $v \in S_{L'}, P(v)$ has a unique selection $P(v) = \{p(v)\}$ where $p(v) = t_v v$ for some $t_v > 0$.*

Proof. For each $v \in S_{L'}$, since $G(t) = J(tv)$ is continuous and satisfies $G(0) = 0$, $G(t) > 0$ when $t > 0$ is small and $G(t) \to -\infty$ as $t \to +\infty$, there exists at least one $t_v > 0$ such that $G(t)$ attains its local maximum at t_v. We have

$$\begin{aligned} \partial G(t) &= t\int_\Omega |\nabla v(x)|^2 dx - \int_\Omega \partial F(tv(x))v(x)dx \\ &= t\left(1 - \int_\Omega \frac{\partial F(tv(x))}{tv(x)} v^2(x)dx\right). \end{aligned}$$

Since $G(t)$ attains its local maximum at $t_v > 0$, we have $0 \in \partial G(t_v)$, or, there exists $\delta F \in \partial F$ such that

$$\int_\Omega \frac{\delta F(t_v v(x))}{t_v v(x)} v^2(x)dx = 1.$$

By the monotonicity Condition (b), for any $t^* > t_v$, it follows $|t^* v(x)| > |t_v v(x)|$ for any $x \in \Omega$ with $v(x) \neq 0$ and then for any selection $\delta^* F \in \partial F$, we have

$$\int_\Omega \frac{\delta^* F(t^* v(x))}{t^* v(x)} v^2(x)dx > \int_\Omega \frac{\delta F(t_v v(x))}{t_v v(x)} v^2(x)dx = 1.$$

Thus such $t_v > 0$ exists and is unique, i.e., we have proved $P(v) = \{p(v)\} = \{t_v v\}$ for each $v \in S_{L'}$. ∎

Next we let $n = 2$ and

$$f(s) = \begin{cases} 4s^3, & \text{if } |s| \le 1, \\ 6s^5, & \text{if } |s| > 1 \end{cases}$$

be a measurable function. Then $F(s) = \int_0^s f(t)dt$ is a Lipschitz continuous function such that

$$F(s) = \begin{cases} s^4, & \text{if } |s| \le 1, \\ s^6, & \text{if } |s| \ge 1, \end{cases} \qquad \partial F(s) = \begin{cases} 4s^3, & \text{if } |s| < 1, \\ [\text{-6, -4}], & \text{if } s = -1, \\ [4,\ 6], & \text{if } s = 1, \\ 6s^5, & \text{if } |s| > 1. \end{cases}$$

It is clear that for any $|s_2| > |s_1| > 0$, and any two selections $\delta_1 F, \delta_2 F \in \partial F$, we have

$$\frac{\delta_2 F(s_2)}{s_2} \in \begin{cases} 4s_2^2, & \text{if } |s_2| < 1, \\ [4,\,6], & \text{if } |s_2| = 1, \\ 6s_2^4, & \text{if } |s_2| > 1. \end{cases} > \frac{\delta_1 F(s_1)}{s_1} \in \begin{cases} 4s_1^2, & \text{if } |s_1| < 1, \\ [4,\,6], & \text{if } |s_1| = 1, \\ 6s_1^4, & \text{if } |s_1| > 1. \end{cases}$$

Therefore both Conditions (a) and (b) are satisfied.

References

[1] A. Ambrosetti and P. H. Rabinowitz, *Dual variational methods in critical point theory and applications*, J.Funct. Anal. 14 (1973), 349-381.

[2] G. Barletta and M. Marano, *Some remarks on critical point theory for locally Lipschitz functions*, Glasow Math. J., 45(2003), 131-141.

[3] H. Brezis and L. Nirenberg, *Remarks on finding critical points*, Commun. Pure Appl. Math. 44 (1991), 939-963.

[4] K. C. Chang, *Variational methods for nondifferentiable functionals and their applications to partial differential equations*, J. Math. Anal. Appl. 80(1981), 102-129.

[5] Frank H. Clarke, Optimization and Nonsmooth Analysis, *Wiley*, New York, 1983.

[6] Z. Denkowski, S. Migorski and N. S. Papageorgiou, An Introduction to Nonlinear Analysis: Theory, *Kluwer Academic Publishers*, Boston, 2003.

[7] W.Y. Ding and W.M. Ni, On the existence of positive entire solutions of a semilinear elliptic equation, *Arch. Rational Mech. Anal.* 91(1986), 238-308.

[8] M. Fabian, P. Habala, P. Hajek, V. M. Santalucia, J. Pelant and V. Zizler, *Functional Analysis and Infinite-Dimensional Geometry*, Springer, New York, 2001.

[9] F. Gazzola and V. Radulescu, *A nonsmooth critical point theory approach to some nonlinear elliptic equations in* $\mathbb{R}^n$, Differential and Integral Equations, 13(2000), 47-60.

[10] N. Halidias, *A nontrivial solution of mountain-pass type for a hemivariational inequality*, Bull. Sci. Math. 127 (2003), no.6, 549–556.

[11] X. Hu, N. Kourogenis and N. S. Papageorgiou, *Nonlinear elliptic eigenvalue problems with discontinuities*, J. Math. Anal. Appl. 233(1999), 406-424.

[12] N. Kourogenis and N. S. Papageorgiou, Nonsmooth critical point theory and nonlinear elliptic equations at resonance, J. Aust. Math. Soc. A 69 (2000), 245-271.

[13] N. Kourogenis, P. Kandilakis and N. S. Papageorgiou, *Two nontrival critical points for nonsmooth functionals applications*, Atti Seminario Matematico Modena, in press.

[14] Y. Li and J. Zhou, *A minimax method for finding multiple critical points and its applications to nonlinear PDEs*, SIAM Sci. Comp., 23(2001), pp.840-865.

[15] Y. Li and J. Zhou, *Convergence results of a local minimax method for finding multiple critical points*, SIAM Sci. Comp., 24(2002), 865-885.

[16] D. Motreanu and P.D. Panagiotopoulos, *Minimax Theorems and Qualitive Properties of the Solutions of Hemivariational Inequalities*, Kluwer Academic Publishers, Boston, 1999.

[17] M. Schechter, Linking Methods in Critical Point Theory, *Birkhauser*, Boston, 1999.

[18] M. Squassina, Existence, Multiplicity and Perturbation Results for Quasilinear Elliptic Problems via Nonsmooth Critical Point Theory, Ph.D. Thesis, Dottorato di Ricerca in Matematica, Universita degli Studi di Milano, Italy, 2002.

[19] 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., to appear.

[20] X. Yao and J. Zhou, *A unified convergence result on a minimax algorithm for finding multiple critical points in Banach spaces*, submitted.

[21] X. Yao and J. Zhou, *Computational theory and methods for solving nonlinear eigenpair problem. Part I: Iso-homogeneous cases*, submitted.

[22] E. Zeidler, Nonlinear Functional Analysis and its Applications III, *Springer-Verlag*, New York, 1985.

[23] J. Zhou, *Instability Analysis of Saddle Points by A Local Minimax Method*, Math. Comput., to appear.