A Minimax Method for Finding Multiple Critical Points and Its Applications to Semilinear PDE

YONGXIN LI* and JIANXIN ZHOU †

Abstract

Most minimax theorems in critical point theory require one to solve a two-level *global* optimization problem and therefore are not for algorithm implementation. The objective of this research is to develop numerical algorithms and corresponding mathematical theory for finding multiple saddle points in a stable way. In this paper, inspired by the numerical works of Choi-McKenna and Ding-Costa-Chen, and the idea to define a solution submanifold, some local minimax theorems are established, which require to solve only a two-level *local* optimization problem. Based on the local theory, a new local numerical minimax method for finding multiple saddle points is developed. The local theory is applied and the numerical method is implemented successfully to solve a class of semilinear elliptic boundary value problems for multiple solutions on some non-convex, non star-shaped and multi-connected domains. Numerical solutions are illustrated by their graphics for visualization. In a subsequent paper [20], we establish some convergence results for the algorithm.

Keywords. Multiple saddle point, Morse index, Local minimax, Semilinear PDE

AMS(MOS) subject classifications. 58E05, 58E30,35A40,35A65

Abbreviated titles. A Local Minimax Method for Multiple Critical Points

1   Introduction

Multiple solutions with different performance and instability indices exist in many nonlinear problems in natural and social sciences. [33, 30, 24, 36, 23]. When cases are variational, the problems can be reduced to solving the Euler-Lagrange equation

$$(1.1) \qquad\qquad\qquad J'(u) = 0,$$

where J, called *a generic energy functional*, is a C^1-functional on a Banach space H and J' or ∇J its Frechet derivative. A solution to the Euler-Lagrange equation (1.1) is called *a critical point* of J. The first candidates for critical points are the local maxima and minima to which the classical critical point theory was devoted

---

in calculus of variation. Traditional numerical methods focus on finding such stable solutions. Critical points that are not local extrema are called *saddle points*, that is, critical points u^* of J, for which any neighborhood of u^* in H contains points v, w s.t. $J(v) < J(u^*) < J(w)$. In physical systems, saddle points appear as *unstable* equilibria or transient excited states. Note that this definition is different from and much more general than the saddle point in optimization and game theory in which a splitting structure for the space H is required to be known in advance and which is therefore not used in critical point theory.

A number $c \in \mathbb{R}$ is a *critical value* of J if $J(\hat{u}) = c$ for some critical point $\hat{u}$. For a critical value c, the set $J^{-1}(c)$ is called a *critical level*. When the second Frechet derivative J'' exists at a critical point $\hat{u}$, then $\hat{u}$ is said to be *nondegenerate* if $J''(\hat{u})$ is invertible. Otherwise $\hat{u}$ is said to be *degenerate*.

Stability is one of the main concerns in control and system design. On the other hand, in many applications, higher maneuverability and performance are desirable, in particular in system design for emergency or combat machineries. Unstable solutions may have much higher maneuverability and performance indices.

Can one find a way to provide a choice or balance between instability and maneuverability or performance indices? Thus one needs to solve for multiple solutions and then study their individual properties.

Numerically finding such unstable solutions in a stable way is very challenging. So far, there is virtually no theory existing in the literature to devise such a feasible numerical algorithm. The objective of this research project is to systematically develop effective numerical algorithms and corresponding mathematical theory for finding multiple saddle points in a stable way. To do so, we need to know local mathematical structure of a critical point and its connection to a critical point at the next critical level. We do not intend to establish *new* existence theorems.

Structure and behavior of critical points have attracted the attention of many researchers. In 1925, Morse proved that if $\hat{u}$ is a nondegenerate critical point of a real function J of n variables, then there exists a neighborhood $\mathcal{N}(\hat{u})$ of $\hat{u}$ and a local homeomorphism h from $\mathcal{N}(\hat{u})$ into H s.t.

$$J(h(u)) = J(\hat{u}) + \frac{1}{2}\langle J''(\hat{u})u, u \rangle \quad \forall u \in \mathcal{N}(\hat{u}).$$

That is to say, J behaves locally like a quadratic function around a nondegenerate critical point. This result, called Morse Lemma, has been extended to a real-valued infinite-dimensional functional ([10]). Therefore critical levels with a local minimum, if it exists, at the bottom can be imagined.

The *Morse Index* (MI) of a critical point $\hat{u}$ of a real-valued functional J is the maximal dimension of a subspace of H on which the operator $J''(\hat{u})$ is negative definite; the *nullity* of a critical point $\hat{u}$ is the dimension of the null-space of $J''(\hat{u})$. Thus for a nondegenerate critical point, if its MI = 0, then it is a local minimizer and a stable solution, and if its $MI > 0$, then it is a saddle point, an unstable solution.

DEFINITION 1.1. *A point $v \in H$ is called a* descent *(ascent) direction of J at a critical point $\hat{u}$, if there exists $\delta > 0$ s.t.*

$$J(\hat{u} + tv) < (>) J(\hat{u}) \quad \forall 0 < |t| < \delta.$$

Thus J has at least k linearly independent descent directions at a critical point with MI $= k$.

Many boundary value problems (BVP) are equivalent to solving [33]

$$(1.2) \qquad A(u) = 0$$

for a solution $u \in H$ and an operator $A : H \to H^*$. When the problem is variational, there exists $J : H \to \mathbb{R}$ s.t.

$$(1.3) \qquad \langle A(u), v \rangle = \langle J'(u), v \rangle = \lim_{t \to 0} \frac{J(u + tv) - J(u)}{t} \qquad \forall v \in H,$$

or $A(u) = J'(u)$. Thus $\hat{u}$ is a (weak) solution to (1.2) if and only if $\hat{u}$ is a critical point of J.

The following semilinear BVP is our model problem in this paper; it is known that this model has originated from many applications in physics, engineering, biology, ecology, geometry, etc. Consider

$$(1.4) \qquad \Delta u(x) - \ell u(x) + f(x, u(x)) = 0 \quad x \in \Omega$$

for $u \in H$ with either the zero Dirichlet boundary condition (B.C.) or the zero Neumann B.C., where Ω is a bounded open domain in $\mathbb{R}^N$, and f is a nonlinear function of $(x, u(x))$ with $u \in H$. For the zero Dirichlet B.C., we let $H = H_0^1(\Omega)$ and $\ell \geq 0$; for the zero Neumann B.C., we let $H = H^1(\Omega)$ and $\ell > 0$, where $H^1(\Omega)$ is the Sobolev space $W^{1,2}(\Omega)$ and $H_0^1(\Omega) = \{ v \in H^1 : v(x) = 0, x \in \partial\Omega \}$ [1]. The associated variational functional is the energy

$$(1.5) \qquad J(u) = \int_\Omega \{ \frac{1}{2} |\nabla u(x)|^2 + \frac{1}{2} \ell u^2(x) - F(x, u(x)) \} \, dx,$$

where $F(x, t) = \int_0^t f(x, \tau) d\tau$ satisfies the assumptions (h1) - (h5) as stated in Section 4 and $u \in H$. Then a direct computation shows that a point $\hat{u} \in H$ is a critical point of J in H if and only if $\hat{u}$ is a weak solution to the BVP (1.4) and, therefore, a classical solution to (1.4) by a standard elliptic regularity argument.

Since Ljusternik-Shnirelman (1934), under a deformation assumption, proved the existence of a saddle point as a minimax solution, i.e., a solution to a two-level optimization problem
$$(1.6) \qquad \min_{A \in \mathcal{A}} \max_{v \in A} J(v)$$

for some collection $\mathcal{A}$ of subsets A in H, minimax principle becomes the most popular approach in critical point theory. Note that there is another minimax approach in multi-level optimization and game theory, which is a two-level optimization of the form
$$\min_{x \in X} \max_{y \in Y} J(x, y)$$

where $H = X \times Y$ for some subspaces X and Y. Due to its splitting structure, this minimax approach prevents from turning-around in a search for a critical point. Although it is known that if $J''(\hat{u})$ is self-adjoint and Fredholm, H has such a splitting

structure around a nondegenerate critical point $\hat{u}$ according to the Morse theory. But such a splitting structure depends on $\hat{u}$ and is not known until one finds $\hat{u}$. Thus this minimax approach does not help in searching for a critical point and is not used in critical point theory.

It is the Mountain Pass Lemma proved in 1973 by Ambrosetti-Rabinowitz [3] by constructing a deformation, that sets a milestone in nonlinear analysis. Since then minimax theorems have gained great popularity in the study of nonlinear PDE and dynamic systems. Various minimax theorems, such as linking and saddle point theorems, have been successfully established to prove existence of multiple solutions to various nonlinear PDE's and dynamic systems [5,8,10,13,14,16, 22,23,24,27,30,31,33, 34,36]. For example, for a semilinear elliptic equation $\Delta u + f(u) = 0$ with the zero Dirichlet B.C., when $f(u)$ is superlinear, it is known that if f is odd, then infinitely many solutions exist; otherwise a third sign-changing solution (MI = 2) exists [34, 5, 9] in addition to a positive and a negative solutions.

When multiple solutions exist in a nonlinear system, some of which are stable and the others are unstable. A stable solution (MI= 0) can be found through local minimization techniques or a monotone iterative scheme (see, e.g., [7], [11], [15], [18], [19], [28] and [29]); However, relatively little is known in the literature on constructing algorithms to compute such unstable saddle points in a numerically stable way. One might mention Newton's method. Note that zeros and critical points are different concepts. Without knowing or using the local structure of a (degenerate) critical point, the usual Newton's method will not be effective or stable. When a local minimization is involved in a quasi-Newton method, it will lead to a local minimum, in case of (1.4), the zero.

Most minimax theorems in the literature mainly focus on the existence issue. They require one to solve a two-level *global* optimization problem, i.e., (constrained) global maximizations at the first level and a global minimization at the second level, and therefore are not for algorithm implementation.

By studying the mountain pass lemma and using an idea from Aubin-Ekeland [4], in 1993 Choi-McKenna [12] proposed a numerical minimax algorithm, called a mountain pass method, to solve the model problem basically for a solution with MI = 1. The algorithm opens a brand new door to numerically compute unstable solutions. The algorithm has been modified in [17] and further revised in [11]. Since the function J in [12] has only one maximum along each direction, whether or not it is a local or global maximum at the first level is not a concern there. The merit of this algorithm is that (a) at the first level, a maximization is taken over an affine line starting from 0 and, (b) a steepest descent direction is used to search for a local minimum at the second level. In contrast, the mountain pass lemma requires a maximization on every continuous path connecting 0 and a given point v at the first level and then a global minimization at the second level. Thus the method in [12] can not be justified by the mountain pass lemma. An earlier result of Ding-Ni [27], which proved for the model problem the existence of a saddle point as a minimax solution that requires a unique maximization on each affine line starting from 0 at the first level and a global minimization at the second level, can only be viewed as a partial justification.

In [11], for a class of functionals, a very simple scheme, called a scaling iterative method is designed to find a solution with MI= 1. That approach is not based on functional analysis. A partial justification of that algorithm is also given there.

A high linking theorem for the existence of a third solution (MI = 2) is proved in [34] by constructing a "local link" at a mountain pass solution. Motivated by this idea, a numerical high linking method is proposed in [17] by Ding-Costa-Chen to solve the model problem for a sign-changing solution (MI = 2). The basic idea is that, assume a mountain pass solution w_1 has been found, by using an ascent direction and a descent direction at w_1, one can form a triangle as a "local linking". Then one can proceed to find a maximum on this triangle. If the maximum is inside the triangle, go to the next step, otherwise, deform the triangle so as to contain this point as an interior point and continue to search for an interior maximum. This method uses constrained maximizations at the first level and a local minimization at the second level. Since in the original version of the high linking theorem, the argument never left a mountain pass solution, and a global minimization is required at the second level, the theorem itself can not serve as a justification of that algorithm. Accordingly, the theoretical justification of that algorithm will be very difficult. The problem is that once an iterate has departed a mountain pass solution, the triangle is no longer a "local linking"; the triangle may degenerate. However, this is the first time in the literature that the idea of a "local linking" is used to find solutions basically with MI = 2.

Inspired by the above numerical works, we developed a new minimax method for finding multiple saddle points. Many numerical results including those presented in Section 5 were obtained in the summer of 1997 and appeared to be very promising. Then we decided to try to establish some mathematical justification for the algorithm and to find out why and under what conditions the algorithm works. This may eventually help us improve the algorithm. This has been soon proved to be a much more challenging task. Since we try to set a general framework and same time always keep some model problems in mind, every time a new model problem is solved numerically, new interesting mathematical questions can be asked. We try to answer a question in a more general way.

For the purpose of mathematical justification, let us adopt another approach. In studying a dynamic system, Nehari [25] introduced the concept of a solution submanifold $\mathcal{M}$, and proved that a global minimizer of the energy functional on $\mathcal{M}$ is a solution to the dynamic system with MI = 1. Ding-Ni used Nehari's idea to study the model problem with the zero Dirichlet B.C. and $\ell = 0$. They defined a solution submanifold

$$(1.7) \qquad \mathcal{M} = \left\{ v \in H_0^1(\Omega) | v \neq 0, \int_\Omega \left[|\nabla v|^2 - v f(v) \right] \, dx = 0 \right\}.$$

Under the condition that $f'(t) > \frac{f(t)}{t}, t \neq 0$, Ding-Ni [27] proved that a global minimizer of the energy function J on $\mathcal{M}$ is a solution with MI = 1 to the model problem.

Our basic idea to design an algorithm for finding multiple saddle points consists of three main elements:

(1) First define a solution (stable) submanifold $\mathcal{M}$ s.t. a local minimum point of $J(u)$ on $\mathcal{M}$ yields a critical point. Thus the problem becomes a minimization of J on the submanifold $\mathcal{M}$ and a saddle point becomes stable on the submanifold $\mathcal{M}$. If a monotone decreasing search is used in the minimization process, the algorithm will be stable. At a point on $\mathcal{M}$, we can apply, e.g., a steepest descent search to approximate a local minimizer of J on $\mathcal{M}$.

(2) There must be a *return rule*. As a steepest descent search usually leaves the submanifold $\mathcal{M}$, for the algorithm to continue to iterate, we need to design a return rule for the search to return to $\mathcal{M}$.

(3) There must be a strategy to *avoid degeneracy*. Since we are searching for a saddle point at a higher critical level, at least, for a new solution, a simple minimization may cause degeneracy to a (old) saddle point at a lower critical level. Thus a strategy to avoid degeneracy is crucial to guarantee that the new critical point found is different from the old ones. This strategy may also be incorporated in the definition of the solution submanifold. It can be seen that:

(1) Choi-McKenna's algorithm has a return rule and a strategy to avoid a degeneracy to the zero with MI $= 0$. We will provide a mathematical justification for their algorithm in Section 2, see Theorem 2.1 with $L = \{0\}$ or Theorem 4.2;

(2) Ding-Costa-Chen's algorithm has a return rule and a strategy to avoid a degeneracy to a solution with MI $= 1$. We will modify their algorithm and then provide a mathematical justification, see Theorem 2.1 with $L = \{w_1\}$.

The organization of this paper is as follows. In Section 2, we first prove an important technical lemma. Then we establish a local minimax characterization of a saddle point, which can be used to design a numerical minimax algorithm for finding multiple saddle points. An existence theorem is also proved in this section. Section 3 is to present our numerical minimax algorithm in detail and some convergence related properties. In Section 4, we apply our minimax method to solve a class of semilinear elliptic equations and present some related analysis. Finally, in Section 5, we exhibit some numerical examples for multiple solutions and their graphics to illustrate the algorithm and theory.

We do not intend to prove any new existence theorem. Our objective is to develop numerical algorithms and corresponding mathematical theory for finding multiple critical points. It is understood that many critical points can not be approximated. Only those with "nice properties" can be numerically approximated. We try to classify those "nice" saddle points through mathematical analysis. It is reasonable that the hypotheses in our local minimax characterization of saddle points in this paper are stronger than that in those existence theorems. Our hypotheses will be gradually localized and generalized as research advances. Methods to check those hypotheses will also be developed in subsequent papers.

2   Local Min-Max Theorems

Let H be a Hilbert space with inner product $\langle , \rangle$ and norm $\| \ \|$, and J be a real C^1-generic energy functional on H. For any subspace $H' \subset H$, denote $S_{H'} =$

$\{v|v \in H', \|v\| = 1\}$ the unit sphere in H'. Let L be a closed subspace in H, called a base subspace, and $H = L \oplus L^{\perp}$ be the orthogonal decomposition where $L^{\perp}$ is the orthogonal complement of L in H. For each $v \in S_{L^{\perp}}$ let $[L, v] = \{tv + w | w \in L, \ t \geq 0\}$ be the closed half subspace. L can be either finite or infinite dimensional.

DEFINITION 2.1. *A set-valued mapping* $P: S_{L^{\perp}} \to 2^H$ *is called the* peak mapping *of* J *w.r.t.* $H = L \oplus L^{\perp}$ *if for any* $v \in S_{L^{\perp}}$, $P(v)$ *is the set of all local maximum points of* J *in* $[L, v]$. *A single-valued mapping* $p: S_{L^{\perp}} \to H$ *is a* peak selection *of* J *w.r.t.* L *if*

$$p(v) \in P(v) \quad \forall v \in S_{L^{\perp}}.$$

For a given $v \in S_{L^{\perp}}$, *we say that* J *has a local peak selection w.r.t.* L *at* v *if there is a neighborhood* $\mathcal{N}(v)$ *of* v *and a function* $p: \mathcal{N}(v) \cap S_{L^{\perp}} \to H$ *s.t.*

$$p(u) \in P(u) \quad \forall u \in \mathcal{N}(v) \cap S_{L^{\perp}}.$$

A special case when $L = \{0\}$, the peak mapping $P(v)$ is the set of all local maximum points of J along the direction v, for any point $v \in S$.

Most minimax theorems in critical point theory used a (constrained) global maximization (on a compact set) at the first level. Thus a solution at the first level always exists. For algorithm implementation, existence is not enough, we want an approximation scheme to a solution. Therefore we use unconstrained local maximization at the first level. Numerically this is great. However, it then raises three major problems in analysis: (a) for some $v \in S_{L^{\perp}}$, $P(v)$ may contain multiple local maxima in $[L, v]$. In particular, P may contain multiple branches, even U-turn or bifurcation points; (b) p may not be defined at some points in $S_{L^{\perp}}$; (c) the limit of a sequence of local maximum points may not be a local maximum point. Thus the analysis involved becomes much more complicated. Although it is well known that the energy function (1.5) of our model problem goes to negative infinity uniformly in any finitely dimensional subspace (See [30]). Thus in any finite-dimensional subspace, there must be at least one maximum point, i.e., $P(v) \neq \emptyset$. That is, Problem (b) will not happen. Problem (a) is not concerned in the numerical works of Choi-McKenna and Ding-Costa-Chen. However for more general settings, all those three problems have to be resolved. As for (a), we use a selection p to choose one branch from others. Numerically it is done by following certain negative gradient flow and developing some consistent strategies to avoid jumps between different branches. For (b), we only need p to exist locally around a point v along a negative gradient flow. We are currently working on an min-orthogonal algorithm where $p(v)$ need only be a local sub-orthogonal point. This will further resolve Problem (b). When P contains some U-turn or bifurcation points, if a saddle point happens to be a U-turn or bifurcation point of P, a local minimax search at either one of the branches connecting to the point will follow the negative gradient flow to the saddle point. If a saddle point is not a U-turn or bifurcation point of P, when a local minimax search is at the same branch of P as the saddle point, according to our local characterization results, it is not a problem; In case when a local minimax search is at a different branch of P connecting to the U-turn or bifurcation point, we will develop a technique to allow the search to pass through a U-turn or bifurcation point. We will address this technique

in a future article. As for Problem (c), more analysis is required. One of the reasons for Problem (c) to take place is that under the definition of a peak selection $p(v)$, the solution submanifold $\mathcal{M} = \{p(v) : v \in S_{L^\perp}\}$ is not closed. In a future paper, we will define a more general local (orthogonal) selection $p(v)$ with which the new solution submanifold is closed and contains the current solution submanifold as a subset. Then we prove that a local minimum point of J on the new solution submanifold also yields a saddle point and the implicit function theorem can be used to check if p is continuous (differentiable) at a given point v, a condition required by all the results we proved in this paper. This study leads to a complete different approach and will be addressed in a future paper.

The following technical lemma plays a crucial role in this paper. It describes the relation between the gradient of J and the variation of a peak selection.

LEMMA 2.1. *For* $v_\delta \in S_{L^\perp}$, *if there is a local peak selection* p *of* J *w.r.t.* L *at* v_δ *s.t. (i)* p *is continuous at* v_δ, *(ii)* $d(p(v_\delta), L) > \alpha > 0$ *and (iii)* $\|\nabla J(p(v_\delta))\| > \delta > 0$, *then there exists* $\varepsilon > 0$, *s.t.*

$$J(p(v(s))) - J(p(v_\delta)) < -\alpha\delta\|v(s) - v_\delta\| \forall 0 < s < \varepsilon$$

and

$$v(s) = \frac{v_\delta + sw}{\|v_\delta + sw\|}, \quad w = -\frac{\nabla J(p(v_\delta))}{\|\nabla J(p(v_\delta))\|}.$$

Proof.   Let $\mathcal{N}(v_\delta)$ be a neighborhood of v_δ for the local peak selection p. By (iii),

$$\langle \nabla J(p(v_\delta)), w \rangle < -\delta.$$

Since $p(v_\delta) \in [L, v_\delta]$, $p(v_\delta) = x_\delta + t_\delta v_\delta$, where $x_\delta \in L$ and $t_\delta \in \mathbb{R}$, we have $t_\delta > \alpha$. $p(v_\delta)$ is a local maximum point of $J(p(x))$ on $[L, v_\delta]$. so it is clear that $w \in L^\perp$ and $w \perp v_\delta$. Due to the continuity, there exist positive numbers ε_1, ε_2, t_1 and t_2 with $0 < t_1 < t_\delta < t_2$ and open balls $B_{\varepsilon_1, L} = \{x | x \in L, \|x - x_\delta\| < \varepsilon_1\}$ and $B_{\varepsilon_2, L^\perp} = \{x | x \in L^\perp, \|x - t_\delta v_\delta\| < \varepsilon_2\}$ s.t.

a) $p(v_\delta)$ is a maximum point of J on $[t_1, t_2] \times v_\delta + B_{\varepsilon_1, L}$,

b) $\langle \nabla J(x_1 + x_2), w \rangle < -\delta, \quad \forall x_1 \in B_{\varepsilon_1, L}, x_2 \in B_{\varepsilon_2, L^\perp}$,

c) $tv_\delta + sw \in B_{\varepsilon_2, L^\perp}$ for any $t \in [t_1, t_2]$ and s, $0 < s < \frac{\varepsilon_2}{2}$.

For any $x_1 \in B_{\varepsilon_1, L}$, and $t \in [t_1, t_2]$, by the mean value theorem,

$$(2.1) \qquad J(tv_\delta + x_1 + sw) - J(tv_\delta + x_1) = \langle \nabla J(tv_\delta + x_1 + \xi w), sw \rangle < -s\delta.$$

By a), $J(tv_\delta + x_1) \leq J(p(v_\delta))$, so we have $J(tv_\delta + x_1 + sw) - J(p(v_\delta)) < -s\delta$. Since $w \perp v_\delta$,

$$(2.2) \qquad \lim_{s \to 0^+} \|\frac{1}{s}(\frac{t_\delta v_\delta + sw}{\|t_\delta v_\delta + sw\|} - v_\delta)\| = \frac{1}{t_\delta} < \frac{1}{\alpha}.$$

Thus without loss of generality, we can assume that for $s > 0$ small

$$(2.3) \qquad \|\frac{t_\delta v_\delta + sw}{\|t_\delta v_\delta + sw\|} - v_\delta\| < \frac{s}{\alpha}.$$

By our notation

$$v(\frac{s}{t_\delta}) = \frac{t_\delta v_\delta + sw}{\|t_\delta v_\delta + sw\|},$$

so we have, for sufficient small s,

(2.4)
$$\alpha \|v(\frac{s}{t_\delta}) - v_\delta\| < s.$$

Combine (2.1) and (2.4), we can find $\varepsilon_3 > 0$, s.t.

$$J(tv_\delta + x_1 + sw) - J(p(v_\delta)) < -\alpha\delta \|v(\frac{s}{t_\delta}) - v_\delta\|, \quad \forall t \in [t_1, t_2], \ x_1 \in B_{\varepsilon_1, L}, \ 0 < s < \varepsilon_3.$$

Denote $D = \{tv_\delta + x_1 + sw | s < \varepsilon_3, \ t \in (t_1, t_2), \ x_1 \in B_{\varepsilon_1, L}\}$. Then D is an open neighborhood of $p(v_\delta) = t_\delta v_\delta + x_\delta$ in the subspace spanned by L, v_δ and w. By the continuity of p at v_δ, there exists positive ε, s.t. $v(\frac{s}{t_\delta}) \in S_{L^\perp} \cap \mathcal{N}(v_\delta)$ and $p(v(\frac{s}{t_\delta})) \in D, \forall \frac{s}{t_\delta} < \varepsilon$. Besides, since $p(v(\frac{s}{t_\delta})) \in [L, v(\frac{s}{t_\delta})]$, $p(v(\frac{s}{t_\delta})) = ct_\delta v_\delta + x_1 + csw$ uniquely for some constant number c and $x_1 \in B_{\varepsilon_1, L}$ with $t_1 < ct_\delta < t_2$ and $cs < \varepsilon_3$. Thus

$$
\begin{aligned}
J(p(v(\frac{s}{t_\delta}))) - J(p(v_\delta)) &= J(ct_\delta v_\delta + x_1 + csw) - J(p(v_\delta)) \\
&< -\alpha\delta\|v(\frac{cs}{ct_\delta}) - v_\delta\| = -\alpha\delta\|v(\frac{s}{t_\delta}) - v_\delta\|. \qquad \blacksquare
\end{aligned}
$$

The following theorem characterizes a saddle point as a local minimax solution and serves as a mathematical justification for our algorithm to be proposed.

THEOREM 2.1. *Let $v_0 \in S_{L^\perp}$. If J has a local peak selection p w.r.t. L at v_0 s.t. (i) p is continuous at v_0, (ii) $d(p(v_0), L) > 0$ and (iii) v_0 is a local minimum point of $J(p(v))$ on $S_{L^\perp}$, then $p(v_0)$ is a critical point of J.*

Proof. Suppose that $p(v_0)$ is not a critical point of J. By Lemma 1, for a sufficient small positive s, set

(2.5)
$$w = -\frac{\nabla J(p(v_0))}{\|\nabla J(p(v_0))\|}, \quad v(s) = \frac{v_\delta + sw}{\|v_\delta + sw\|},$$

we have

$$J(p(v(s))) - J(p(v_0)) \le -\frac{1}{2}\|\nabla J(p(v_0))\|\|v(s) - v_0\|d(p(v_0), L) < 0$$

It contradicts the fact that v_0 is a local minimum point of $J(p(v))$ on $S_{L^\perp}$. Thus $p(v_0)$ must be a critical point of J. $\blacksquare$

REMARK 2.1. Condition (i) is required for a numerical algorithm to be stable and convergent. Condition (ii) is a separation condition posed by almost all minimax theorems to have a new critical point. Condition (iii) replaces a global minimization, used in most minimax theorems in the literature, by a local minimization. Due to the local nature of the above characterization and (2.5), it is clear that $S_{L^\perp}$ in Condition (iii) can be localized to be any subset containing $v(s)$ for small $s \ge 0$.

To establish an existence result, the following PS condition is used to replace the usual compactness condition.

DEFINITION 2.2. *A function $J \in C^1(H)$ is said to satisfy the Palais-Smale (PS) condition, if any sequence $\{u_n\} \in H$ with $J(u_n)$ bounded and $J'(u_n) \to 0$ has a convergent subsequence.*

THEOREM 2.2. *Let J be C^1 and satisfy (PS) condition. If there is a peak selection p of J w.r.t. L, s.t. (i) p is continuous, (ii) $d(p(v), L) \geq \alpha$, for some positive α and for any $v \in S_{L^\perp}$, (iii)] $\inf_{v \in S_{L^\perp}} J(p(v)) > -\infty$, then there exists $v_0 \in S_{L^\perp}$ s.t. $p(v_0)$ is a critical point of J, and*

$$J(p(v_0)) = \min_{v \in S_{L^\perp}} J(p(v)).$$

Proof. Since $S_{L^\perp}$ is a closed metrical subspace, $J(p(v))$ is a continuous function on $S_{L^\perp}$, bounded from below, by Ekeland's variational principle, for any integer n, there exists $v_n \in S_{L^\perp}$ s.t.

$$(2.6) \qquad J(p(v_n)) \leq \inf_{v \in S_{L^\perp}} J(p(v)) + \frac{1}{n}$$

and for any $v \in S_{L^\perp}, v \neq v_n$,

$$J(p(v)) - J(p(v_n)) \geq -\frac{1}{n}\|v - v_n\|.$$

From Lemma 2.1, for some v in $S_{L^\perp}$ and close to v_n,

$$J(p(v)) - J(p(v_n)) \leq -\frac{1}{2}\|v - v_n\|\|\nabla J(p(v_n))\|d(p(v_n), L).$$

Thus

$$(2.7) \qquad \|\nabla J(p(v_n))\| \leq \frac{2}{nd(p(v_n), L)} < \frac{2}{n\alpha}.$$

By (PS) condition, $\{p(v_n)\}$ has a subsequence, denoted again by $\{p(v_n)\}$, converging to some point $u_0 \in H$. Note that $p(v_n) = t_n v_n + x_n$ for some scalar $t_n > 0$, $v_n \in S_{L^\perp}$ and $x_n \in L$. It follows from $\|p(v_n) - p(v_m)\|^2 = \|t_n v_n - t_m v_m\|^2 + \|x_n - x_m\|^2$ that $\{t_n v_n\}$ is a Cauchy sequence as well. Since $\|v_n\| = 1$, $\|t_n v_n\| = t_n \to t_0$. By our assumption (ii), $t_0 \geq \alpha > 0$. Thus $v_n \to v_0 \in S_{L^\perp}$. By the continuity, we have $u_0 = p(v_0)$. Then by (2.7), $p(v_0)$ is a critical point of J and moreover, (2.6) leads to

$$J(p(v_0)) = \min_{v \in S_{L^\perp}} J(p(v)). \qquad \blacksquare$$

The following corollary is a special case of Theorem 2.2 with $L = \{0\}$, which can be viewed as a mathematical justification of the modified Choi-McKenna's algorithm to find a solution with MI$= 1$.

COROLLARY 2.1. *Let $J \in C^1(H)$ and satisfy (PS) condition. Let S be the unit sphere of H and $p(v)$ be a local maximum point of J on $\{tv | t \in (0, +\infty)\}$ for each $v \in S$ s.t. (i) p is continuous, (ii) $\|p(v)\| \geq \alpha$, for some $\alpha > 0$ and for any $v \in S$, (iii) $\inf_{v \in S} J(p(v)) > -\infty$, then there exists $v_0 \in S$ s.t. $p(v_0)$ is a critical point of J, and moreover,*

$$J(p(v_0)) = \min_{v \in S} J(p(v)).$$

Proof. We can simply apply Theorem 2.2 with $L = \{0\}$ to draw the conclusion. ∎

REMARK 2.2. To apply Ekeland's variational principle, a function must be bounded from below. In general, J is not bounded from below. However $J(p(v))$ has a much better chance to be bounded from below. Therefore we use Condition (iii) in Theorem 2.2. This condition is satisfied automatically by our model problem, see Section 4.

3   A Minimax Algorithm For Finding Critical Points

If we define a solution submanifold by

$$\mathcal{M} = \{p(v) : v \in S_{L^\perp}\},$$

then by Theorem 2.1, a local minimum point $p(v_0)$ of J on $\mathcal{M}$ is a critical point of J. Numerically, a local minimum point can be approximated by the steepest descent search. When the steepest descent search leaves $\mathcal{M}$ to a point v, a local maximization on $[L, v]$ yields $p(v)$, a point returned to $\mathcal{M}$. Thus our algorithm can continue to iterate. Since $J'(p(v(s))) \perp L$, the search will not degenerate to L which contains previously found solutions.

The Flow Chart of a Minimax Algorithm

Step 1: Let $n-1$ critical points $w_1, w_2, \ldots, w_{n-1}$ of J be previously found and w_{n-1} have the highest critical value. Set the base space $L = \text{span}\{w_1, w_2, \ldots, w_{n-1}\}$. Let $v^0 \in L^\perp$ be an ascent direction at w_{n-1} and $\bar{\alpha}_k > 0$ be given;

Step 2: For $k = 0$, solve

$$w^k \equiv p(v^0) \equiv t_0^* v^0 + v_L^* = \arg \max_{u \in [L, v^0]} J(u);$$

Step 3: Compute the steepest descent direction d^k of J at w^k;

Step 4: if $\|d^k\| \leq \varepsilon$ then output $w_n \equiv w^k$ and stop, else goto Step 5;

Step 5: For each $0 < \alpha \leq \bar{\alpha}_k$, denote $v^k(\alpha) = \dfrac{v^k - \alpha * d^k}{\|v^k - \alpha * d^k\|}$ and with the initial guess $u = t_0^* v^k(\alpha) + v_L^*$, solve

$$p(v^k(\alpha)) = \arg \max_{u \in [L, v^k(\alpha)]} J(u);$$

then solve

$$w^{k+1} \equiv p(v^{k+1}) \equiv t_0^* v^k(\alpha^*) + v_L^* = \arg \min_{0 < \alpha \leq \bar{\alpha}_k} J(p(v^k(\alpha)));$$

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

REMARK 3.1. Let us make some remarks on each step in the above algorithm.

In Step 1, when we start from some known critical points, if the critical point w_0 with MI = 0 is not zero, we should add w_0 to the base space L. We first start from w_0 to find w_1 and so on. As for our model problem, $w_0 = 0$, so $L = \{0\}$ and according to the Morse theory, to find w_1, we may use any direction $v^0 \in H$ as an ascent direction of J at w_0.

Now assume $w_1, w_2, \ldots, w_{n-1}$ have been found this way and w_{n-1} is the one with the highest critical value. The base space L is spanned by $w_1, w_2, \ldots, w_{n-1}$. We start the algorithm with an ascent direction v^0 of J at w_{n-1}.

It is clear that the separation condition is quite reasonable to ensure that the critical point found is not in the base space L spanned by the old critical points, therefore will be a new one at a higher critical level.

The idea to choose a direction $v^0 \in L^\perp$ as an ascent direction is quite useful in practice. It makes the separation condition easier to satisfy. For example, in Section 4, to choose an ascent direction, we alway use a functional in $H_0^1(\Omega)$ whose peak is away from the peaks of the known critical points $w_1, w_2, \ldots, w_{n-1}$. This idea works very well.

In Step 2, a simple unconstrained local maximization problem.

In Step 3, to find the steepest descent direction d^k of J at a point w^k is usually equivalent to solving a linear system. As for the model problem, a direct calculation shows that the steepest descent direction d^k of J at w^k can be obtained by solving the following linear elliptic BVP

$$(3.1) \quad \begin{cases} \Delta d^k(x) - \ell d^k(x) = \Delta w^k(x) - \ell w^k(x) + f(x, w^k(x)), & x \in \Omega \\ d^k(x) = 0, \quad x \in \partial\Omega & \text{for the zero Dirichlet B.C.,} \\ \dfrac{\partial d^k(x)}{\partial n} = \dfrac{\partial w^k(x)}{\partial n}, \quad x \in \partial\Omega & \text{for the zero Neumann B.C.} \end{cases}$$

which can be solved numerically by various finite element, finite difference or boundary element solvers. Since one important topic in nonlinear analysis is to study how a variation of the domain affects the profile of a solution [14] and the boundary element method can easily handle a complex domain or trace a variation of the domain, we use a boundary element method. Note that for the zero Neumann B.C. problem, if we choose an initial ascent direction v^0 with $\frac{\partial v^0}{\partial n}(x)|_{\partial\Omega} = 0$, then in every iterate, we have $\frac{\partial d^k}{\partial n}(x)|_{\partial\Omega} = 0$ in (3.1).

In Step 4, for our model problem, we can use a norm $\|\Delta d^k - l \cdot d^k\|_{L^2} < \varepsilon$ to control the error. Note that this an absolute error indicator for our model problem, since

$$\Delta d^k(x) - l \cdot d^k(x) = \Delta w^k(x) - \ell w^k(x) + f(x, w^k(x)).$$

In Step 5, for each point $v^k(\alpha) = v^k - \alpha \cdot d^k$ in the steepest descent direction, we can find a local maximum $p(v^k(\alpha))$ of J in $[L, v^k(\alpha)]$. We then try to find the point in the steepest descent direction with the smallest such maximum. We must follow a consistent way to find a local maximum point of J so that $p(v)$ depends on v continuously and is kept away from L. Thus we specify the initial guess $u = t_0^* v^k(\alpha) + v_L^*$ in searching for a local maximum in $[L, v^k(\alpha)]$. This initial

guess closely and consistently traces the position of the previous point $w^k = t_0^* v^k + v_L^*$. This strategy is also to avoid the algorithm from possible oscillating between different branches of the peak mapping P.

The number $\bar{\alpha}$ is to enhance the stability of the algorithm. It controls the step-size of the search along the steepest descent direction to avoid the search to go too far, i.e., to leave the solution (stable) submanifold $\mathcal{M}$ too far to lose stability of the algorithm. ▮

The following theorem implies that the local minimax algorithm is strictly descending and therefore a stable algorithm.

THEOREM 3.1. *If $d^k = \nabla J(w^k) \neq 0$, $w^k \notin L$ and p is continuous at v^k, then*

$$J(w^{k+1}) < J(w^k).$$

If p is continuous and $d(p(v), L) > \alpha > 0$ for all $v \in S_{L^\perp}$, then there exist $\bar{\alpha}_k > 0$ and $d > 0$ s.t.

$$(3.2) \qquad J(w^{k+1}) - J(w^k) < -d\|\nabla J(w^k)\|\|v^{k+1} - v^k\| \quad \forall k = 1, 2, ...,$$

where $w^k = p(v^k)$ and $w^{k+1} = p(v^{k+1})$ are determined in Step 5 of the algorithm.
Proof. We only have to prove (3.2). Using the notation in the algorithm scheme. Write

$$v^{k+1} = v^k(\alpha_k^*) = \frac{v^k - \alpha_k^* d^k}{\|v^k - \alpha_k^* d^k\|}.$$

Take $\delta = \|d^k\|/2$, by Lemma 2.1, we can find $\bar{\alpha}_k > 0$, s.t. if $\bar{\alpha}_k > s > 0$ then

$$J(p(v^k(s))) - J(p(v^k)) < -\alpha\delta\|v^k(s) - v^k\|.$$

In particular, we choose such $\bar{\alpha}_k$ in the algorithm and $0 < \alpha_k^* \leq \alpha_k$ is satisfied, with $d = 2\alpha$, we have

$$J(p(v^k(\alpha_k^*))) - J(p(v^k)) < -d\|d^k\|\|v^k(\alpha) - v^k\|,$$

i.e., (3.2). ▮

Convergence is always a paramount issue for any numerical algorithm. Due to multiplicity, degeneracy and instability of saddle points, general convergence analysis will be very difficult. More profound analysis is required. We will establish some convergence results of the algorithm in a subsequent paper [20]. Instead, in Section 4, we present some applications of our minimax theorem and method to a class of semilinear elliptic PDE.

4   Application to Semilinear Elliptic PDE

In this section, we apply our local minimax method to study the model problem. We use the notations as in [30] with a slight change. Ω is a smooth bounded domain in $\mathbb{R}^n$. Consider a semilinear elliptic Dirichlet BVP

$$(4.1) \qquad \begin{cases} \Delta u(x) + f(x, u(x)) & = & 0, & x \in \Omega, \\ u(x) & = & 0, & x \in \partial\Omega, \end{cases}$$

where the function $f(x, \xi)$ satisfies the following standard hypothesis:

(h1) $f(x,\xi)$ is locally Lipschitz on $\bar{\Omega} \times \mathbb{R}$;

(h2) there are positive constants a_1 and a_2 s.t.

$$(4.2) \qquad |f(x,\xi)| \le a_1 + a_2|\xi|^s$$

where $0 \le s < \frac{n+2}{n-2}$ for $n > 2$. If $n = 2$,

$$(4.3) \qquad |f(x,\xi)| \le a_1 \exp \phi(\xi)$$

where $\phi(\xi)\xi^{-2} \to 0$ as $|\xi| \to \infty$;

(h3) $f(x,\xi) = o(|\xi|)$ as $\xi \to 0$;

(h4) there are constants $\mu > 2$ and $r \ge 0$ s.t. for $|\xi| \ge r$,

$$(4.4) \qquad 0 < \mu F(x,\xi) \le \xi f(x,\xi),$$

where $F(x,\xi) = \int_0^\xi f(x,t)dt$.

In our later numerical computation, we solve problems in $\mathbb{R}^2$, where (h2) is not a substantial restriction. (h4) says that f is superlinear, which implies that there exist positive numbers a_3 and a_4 s.t. for all $x \in \bar{\Omega}$ and $\xi \in \mathbb{R}$

$$(4.5) \qquad F(x,\xi) \ge a_3|\xi|^\mu - a_4.$$

The variational functional associated to the Dirichlet problem (4.1) is

$$(4.6) \qquad J(u) = \frac{1}{2}\int_\Omega |\nabla u(x)|^2 dx - \int_\Omega F(x,u(x))dx, \quad u \in H \equiv H_0^1(\Omega),$$

where we use an equivalent norm $\|u\| = \int_\Omega |\nabla u(x)|^2 dx$ for the Sobolev space $H = H_0^1(\Omega)$.

It is well known [30] that under Conditions (h1) through (h4), J is C^1 and satisfy (PS) condition. A critical point of J is a weak solution, and also a classical solution of (4.1). 0 is a local minimum point (MI$= 0$) of J. Moreover, in any finitely dimensional subspace of H, J goes to negative infinity uniformly. Therefore, for any finite dimensional subspace L, the peak mapping P of J w.r.t. L is nonempty. We need one more hypothesis, that is

(h5) $\frac{f(x,\xi)}{|\xi|}$ is increasing w.r.t. ξ, or

(h5') $f(x,\xi)$ is C^1 w.r.t. ξ and $f_\xi(x,\xi) - \frac{f(x,\xi)}{\xi} > 0$.

It is clear that (h5') implies (h5). If $f(x,\xi)$ is C^1 in ξ, then (h5) and (h5') are equivalent. All the power functions of the form $f(x,\xi) = |\xi|^k\xi$ with $k > 0$, satisfies (h1) through (h5'), and so do all the positive linear combinations of such functions. Under (h5) or (h5'), J has only one local maximum point in any direction, or, the peak mapping P of J w.r.t. $L = \{0\}$ has only one selection. In other words, $P = p$. The proof can be found in [27] and [14].

LEMMA 4.1. *Under (h5) or (h5'), for any $u \in H$, the function $g(t) = J(tu)$, $t \ge 0$, has a unique local and so global maximum point.* ∎

Let $L = \{0\}$ and $\mathcal{M} = \{p(v)|v \in S_{L^\perp} = S_H\}$ where $p(v)$ is the unique peak selection of J w.r.t. L. By the above lemma, it can be easily checked that $\mathcal{M}$ is exactly the solution submanifold (1.7) defined by Ding and Ni. Our definition displays the essence why such a solution submanifold works. Also our definition is given in a more general way. It also works for finding a critical point as a local minimax solution with a higher MI. Actually, under (h5), the peak selection is not only unique but also continuous. In other words, $\mathcal{M}$ is a topological manifold. The following theorem is given in a more general form, which mainly states that uniqueness implies continuity.

THEOREM 4.1. *Under the hypothesis (h1) through (h5), if the peak mapping P of J w.r.t. a finitely dimensional subspace L is singleton at $v_0 \in S_{L^\perp}$ and for any $v \in S_{L^\perp}$ around v_0, a peak selection $p(v)$ is a global maximum point of J in $[L, v]$, then p is continuous at v_0.*

Proof. (h4) implies (4.5), namely, $F(x, \xi) \geq a_3|\xi|^\mu - a_4$, where a_3 and a_4 are positive constants depending on F and Ω. It is known that $\int_\Omega |u(x)|^\mu dx$ is a positive continuous functional in H and $S_H \cap [L, v_0]$ is compact, thus we can write

$$\alpha_0 \equiv \min_{v \in S_H \cap [L, v_0]} \int_\Omega |v(x)|^\mu dx > 0.$$

Let $\alpha = \dfrac{\alpha_0}{2}$. For each $v \in S_H \cap [L, v_0]$ there is a neighborhood $\mathcal{N}(v)$ of v s.t.

$$\int_\Omega |u(x)|^\mu dx > \alpha \quad \forall u \in \mathcal{N}(v).$$

Since

$$\left(S_H \cap [L, v_0]\right) \subset \bigcup_{v \in S_H \cap [L, v_0]} \mathcal{N}(v)$$

and $S_H \cap [L, v_0]$ is compact, there exist $v_1, ..., v_n \in S_H \cap [L, v_0]$ s.t.

$$\left(S_H \cap [L, v_0]\right) \subset \cup_{i=1}^n \mathcal{N}(v_i) \quad \text{and} \quad \int_\Omega |u(x)|^\mu dx > \alpha \quad \forall u \in \cup_{i=1}^n \mathcal{N}(v_i).$$

Note that for each $v \in S_{L^\perp}$ and $w \in S_H \cap [L, v]$, we can write $w = w_l + t_w v$ with $w_L \in L$ and $|t_w| \leq 1$. Then $w_0 = w_l + t_w v_0 \in S_H \cap [L, v_0]$ and $\|w - w_0\|^2 = t_w^2\|v - v_0\|^2 \leq \|v - v_0\|^2$. Thus we can find a neighborhood D of v_0, s.t.

$$\left(S_H \cap [L, v]\right) \subset \cup_{i=1}^n \mathcal{N}(v_i) \quad \forall v \in D \cap S_{L^\perp}.$$

Therefore we have

(4.7) $$\int_\Omega |w|^\mu dx \geq \alpha, \quad \forall v \in D \cap S_{L^\perp}, w \in S_H \cap [L, v].$$

For any $v \in D \cap S_{L^\perp}$, $p(v)$ is a maximum point of J in $[L, v]$. In particular, $p(v)$ is a maximum point of J on the half line $\{tp(v)| t \in \mathbb{R}^+\}$. Set $w = \dfrac{p(v)}{\|p(v)\|}$ and define

$$g(t) = J(tw) = \frac{1}{2}t^2 \int_\Omega |\nabla w(x)|^2 \, dx - \int_\Omega F(x, tw(x)) \, dx.$$

Then g is positive and increasing near 0, and goes to $-\infty$ as $t \to \infty$. A local maximum is solved from

$$0 = \frac{dg}{dt}|_{t=t_1} = t_1 \int_\Omega |\nabla w(x)|^2 \, dx, - \int_\Omega w(x) f(x, t_1 w(x)) \, dx$$

or

(4.8) $$0 = \|p(v)\| \int_\Omega |\nabla w|^2 dx - \int_\Omega w f(x, \|p(v)\|w) dx.$$

Since $\|w\| = 1$, by using (4.4) and (4.7), we have

$$
\begin{aligned}
1 &= \frac{1}{\|p(v)\|^2} \int_\Omega \|p(v)\| w f(x, \|p(v)\|w) dx \geq \frac{1}{\|p(v)\|^2} \int_\Omega \mu F(x, \|p(v)\|w) dx \\
&\geq \frac{1}{\|p(v)\|^2} \int_\Omega \mu(a_3 \|p(v)\|^\mu |w|^\mu - a_4) dx \\
&= a_3 \mu \|p(v)\|^{\mu-2} \int_\Omega |w|^\mu dx - \frac{1}{\|p(v)\|^2} \int_\Omega \mu a_4 dx \\
&\geq \mu a_3 \|p(v)\|^{\mu-2} \alpha - \frac{1}{\|p(v)\|^2} \int_\Omega \mu a_4 dx.
\end{aligned}
$$

Note that $\mu > 2$, the right hand side of the last inequality goes to ∞ if $\|p(v)\| \to \infty$ and this violates the the above inequalities. Therefore, there must exist $\beta > 0$ s.t. $\|p(v)\| \leq \beta$.

Now let $\{v_n\} \subset D \cap S_{L^\perp}$ be any sequence s.t. $v_n \to v_0$. Denote $p(v_n) = t_n v_n + x_n$, where $x_n \in L$. Since $\|p(v_n)\| \leq \beta$ and $v_n \perp x_n$, we have $t_n \leq \beta$ and $\|x_n\| \leq \beta$. Therefore we can find a subsequence $\{v_{n_k}\}$ s.t. t_{n_k} and x_{n_k} converge, respectively, to t_0 and x_0. In other words, $p(v_{n_k})$ goes to $t_0 v_0 + x_0$, which lies in $[L, v_0]$. Since we assume that $p(v_{n_k})$ is a global maximum point of J in $[L, v_{n_k}]$ for each n_k, the limit point $t_0 v_0 + x_0$ must be a maximum point of J in $[L, v_0]$ as well. But by the assumption, the peak mapping P of J is singleton at v_0, so $t_0 v_0 + x_0 = p(v_0)$. Since $\{v_n\}$ is arbitrary, by the above argument, p is continuous at v_0.   ∎

As an immediate conclusion of Theorem 4.1, we have the following continuous result (see Lemma 4.1 in [36]).

COROLLARY 4.1. *Under the hypothesis (h1) through (h5), the only peak selection p of J w.r.t. $L = \{0\}$ is continuous.*
Proof. By Lemma 4.1, there is only one peak selection p of J w.r.t. $L = \{0\}$. In any direction v, the function $g(t) = J(tu)$ possesses only one local maximum point, therefore a global maximum point of J over the subset $\{tv | t \in \mathbb{R}^+\}$. Thus from Theorem 4.1, it is continuous at any point.   ∎

Moreover, the unique selection p of peak mapping w.r.t. $L = \{0\}$ satisfies all the requirements of Theorems 2.1 and 2.2. Thus we can apply Theorems 2.1 and 2.2 to establish the following existence result

THEOREM 4.2. *Under the hypothesis of (h1) through (h5), there exists at least one solution to*

(4.9) $$\text{local} \min_{x \in \mathcal{M}} J(x)$$

and any such a solution is a critical point of J, and therefore a solution to problem (4.1).

Proof. $\mathcal{M}$ is the image of the unique peak selection p of J w.r.t. $L = \{0\}$. By Corollary 4.1, p is continuous. Under the conditions of (h2) through (h4), we know that (see [30]),

$$J(u) = \frac{1}{2}\|u\|^2 + o(\|u\|^2)$$

as $u \to 0$. Thus we can find $\delta > 0$ s.t. $\|p(v)\| > \delta$ for any direction $v \in H$. This is exactly the separation condition in Theorems 2.1 and 2.2. Obviously, $J(p(v)) > 0$ for each direction $v \in H$, thus is bounded from below. Therefore all the conditions in Theorems 2.1 and 2.2 are satisfied. Theorem 2.2 states that there is at least one critical point as a minimax solution and Theorem 2.1 confirms that any local minimax solution is a critical point. ∎

By a similar argument as in the proof of the above theorem and taking Lemma 4.1 into account, we can show that for BVP (4.1), for any closed subspace L and any peak selection p of J w.r.t. L, we have $\inf_{v \in S_{L^\perp}} J(p(v)) > \delta > 0$.

It is known that for BVP (4.1), solution prefers open space. When the domain has multiple compartments connected by narrow corridors, such as a dumbbell-shaped domain in Section 5 for our computational examples, multiple solutions do exist as local minimax solutions. The one with the smallest energy is the global minimax solution, i.e., the ground state.

The following results indicates that under Conditions (h1) to (h5'), for $L = \{0\}$, the minimax algorithm is actually a minimization on a differentiable submanifold $\mathcal{M}$.

THEOREM 4.3. *Assume that Conditions (h1) – (h5') are satisfied and that there exist $a_5 > 0$ and $a_6 > 0$ s.t. for s as specified in (h2),*

$$(4.10) \qquad |f_\xi(x,\xi)| \le a_5 + a_6|\xi|^{s-1}.$$

Then the only peak selection p of J w.r.t. $L = \{0\}$ is C^1.
Proof. Set $G: S_H \times \mathbb{R}^+ \to \mathbb{R}$, $G(v,t) = t - \int_\Omega v(x)f(x,tv(x))dx$. Thus under (4.10), G is C^1. Denote $\mathcal{M} = \{p(v)|v \in S_H\}$ where $p(v)$ is the only peak selection of J w.r.t. $L = \{0\}$, i.e., $p(v)$ is the maximum point of J on $\{tv|t > 0\}$. As in the proof of Theorem 4.1, for any v, if $p(v) = tv$, then we have

$$0 = t - \int_\Omega v(x)f(x,tv(x))dx$$

Thus $\mathcal{M}$ is essentially the inverse image of G at 0, i.e., $\mathcal{M} = G^{-1}(0)$, and $\|p(v)\|$, as a positive number is the solution of $t(v)$ to the equation $G(v,t(v)) = 0$. On the other hand

$$(4.11) \qquad \frac{\partial G}{\partial t} = 1 - \int_\Omega v^2(x)f_\xi(x,tv(x))\,dx.$$

For each v_0, at each pair (v_0,t_0) with $G(v_0,t_0) = 0$,

$$0 = 1 - \int_\Omega v_0^2(x)\frac{f(x,t_0v_0(x))}{t_0v_0(x)}\,dx > 1 - \int_\Omega v_0^2 f_\xi(x,t_0v_0(x))\,dx, \quad \text{(by (p5'))},$$

i.e., $\frac{\partial G}{\partial t} < 0$ at (v_0,t_0), provided $G(v_0,t_0) = 0$. By the implicit function theorem, the solution $t(v)$ to the equation $G(v,t(v)) = 0$ exists uniquely in a neighborhood of each v_0 and is C^1 in v. Therefore $\|p(v)\|$ is a C^1 function in v, and so is $p(v)$. ∎

Actually, in the proof, we can see that the solution submanifold, $\mathcal{M}$, is a differentiable manifold, because Lemma 2.1 implies

$$\|\nabla J|_{\mathcal{M}}\| \geq \delta \|\nabla J\| \quad \text{for some } \delta > 0.$$

EXAMPLE 4.1. *Let us consider the BVP on a smooth bounded domain* $\Omega \subset \mathbb{R}^n$ *for* $p > 2$

$$(4.12) \qquad \begin{cases} \Delta u(x) + |u(x)|^{p-2}u(x) = 0, & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega. \end{cases}$$

The associated variational functional is

$$J(u) = \frac{1}{2}\int_{\Omega}|\nabla u|^2\,dx - \frac{1}{p}\int_{\Omega}|u(x)|^p\,dx, \quad u \in H = H_0^1(\Omega).$$

For each $v \in S$, *let* $u = tv, t > 0$, *then*

$$J(tv) = \frac{t^2}{2}\int_{\Omega}|\nabla v|^2\,dx - \frac{t^p}{p}\int_{\Omega}|v(x)|^p\,dx = \frac{t^2}{2} - \frac{t^p}{p}\int_{\Omega}|v(x)|^p\,dx.$$

Thus

$$0 = \frac{\partial}{\partial t}J(tv) = t - t^{p-1}\int_{\Omega}|v(x)|^p\,dx$$

leads to

$$t_v = \left[\frac{1}{\int_{\Omega}|v(x)|^p\,dx}\right]^{\frac{1}{p-2}} > 0.$$

The peak selection p *of* J *w.r.t.* $L = \{0\}$ *is*

$$p(v) = t_v v = \left[\frac{1}{\int_{\Omega}|v(x)|^p\,dx}\right]^{\frac{1}{p-2}} v, \quad \forall v \in S$$

a continuously differentiable function and the solution manifold

$$\mathcal{M} = \left\{ \left[\frac{1}{\int_{\Omega}|v(x)|^p\,dx}\right]^{\frac{1}{p-2}} v : v \in S \right\}$$

is a differentiable manifold.

5   Computational Examples

We have applied our numerical algorithm to solve many semilinear BVP with zero Dirichlet B.C. on various domains, such as the Lane-Emden equation, the Henon's equation and the Chandrasekhar equation on a disk, rectangle, concentric annulus, nonconcentric annulus, dumbbell-shaped domains and dumbbell-shaped domains with cavities. Here we present the computational results for the Lane-Emden equation

$$(5.1) \qquad \begin{cases} \Delta u(x) + u^3(x) = 0, & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega, \end{cases}$$

where the domain Ω is, respectively, a dumbbell-shaped domain, a dumbbell-shaped domain with cavities (nonsymmetric) and a concentric annulus (highly degenerate). Here the solution $u(x)$ represents the density, so we are interested only in positive solutions. In all examples, we use a *cos* function to create "mound" shaped function as an initial ascent direction v_0 and a norm $\|\Delta \hat{u}\|_{L^2} = \|\Delta u + u^3\|_{L^2} < \varepsilon$ to control the error and terminate the iterate. Some solutions with MI = 1 have been computed elsewhere, see, e.g., [12], [17], [11] and references therein. It is to the best of our knowledge that those solutions with higher MI are the first time to be computed.

Case 1: On a dumbbell-shaped domain.
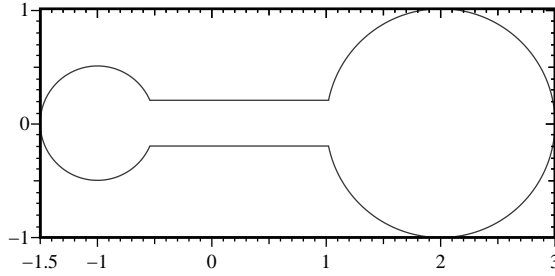
Fig. 1.   *A dumbbell-shaped domain.*

We use, respectively, the following three "mound" functions as initial ascent directions.

$$v_0^i(x) = \begin{cases} \cos(\dfrac{|x - x_i|}{d_i}\dfrac{\pi}{2}) & \text{if } |x - x_i| \leq d_i, \\ 0 & \text{otherwise,} \end{cases}$$

where $x_1 = (2,0), d_1 = 1, x_2 = (-1,0), d_2 = 0.5, x_3 = (0.25, 0), d_3 = 0.2$.
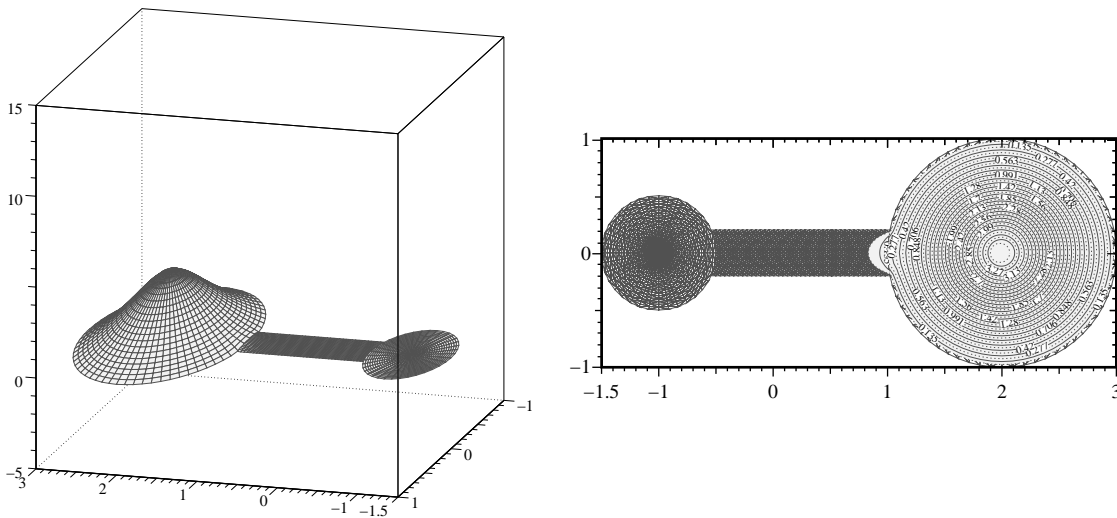
Fig.   2.   *The ground state solution w_1^1 with MI = 1 and its contours. $v_0 = v_0^1$,* $\varepsilon = 10^{-4}, J = 10.90, u_{\max} = 3.652.$
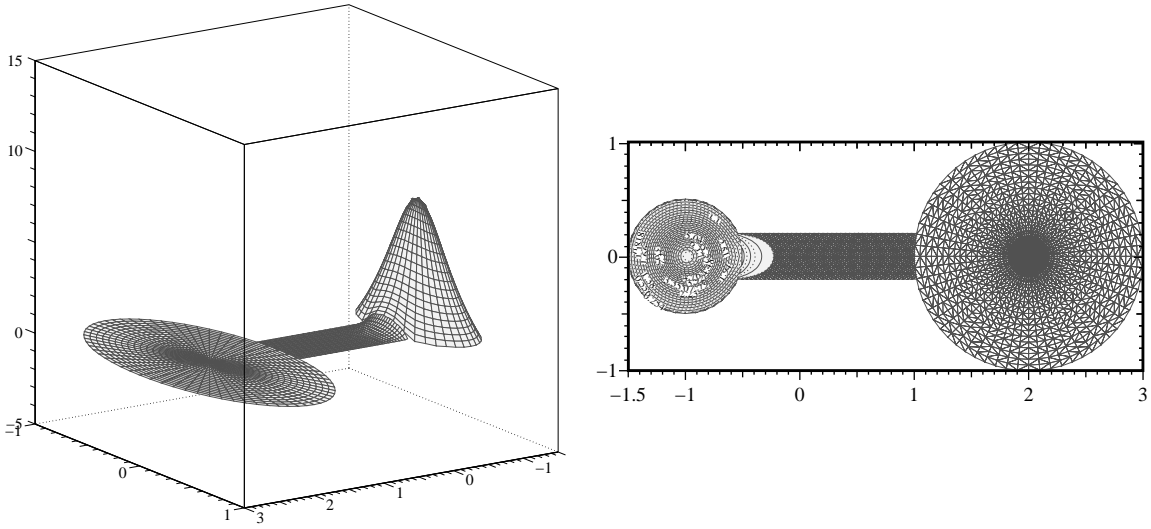
FIG. 3.   *The second solution w_1^2 with $MI = 1$ and its contours. $v_0 = v_0^2, \varepsilon = 10^{-4}$, $J = 42.22, u_{\max} = 7.037$.*
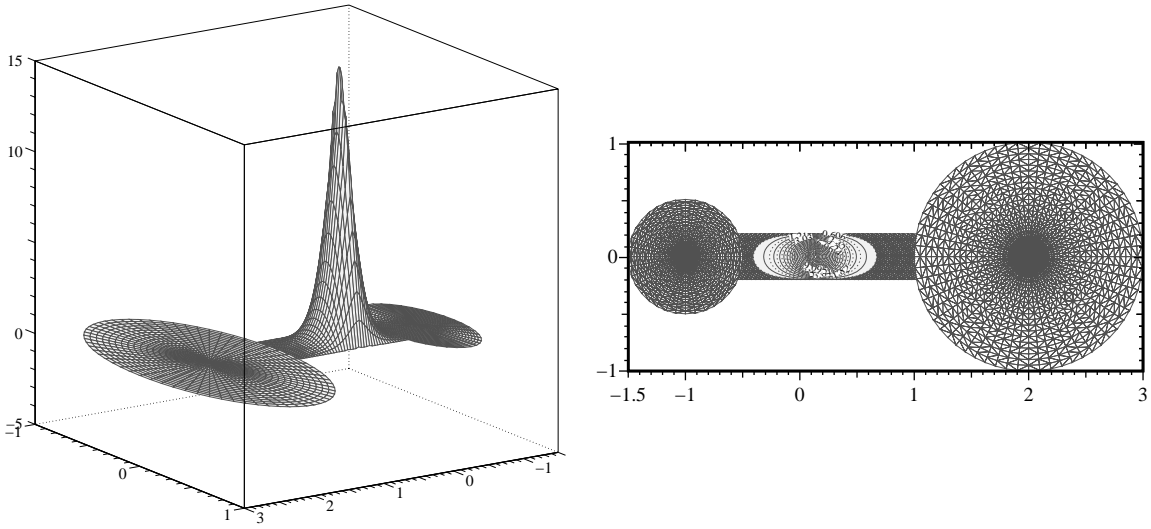
FIG. 4.   *The third solution w_1^3 with $MI = 1$ and its contours. $v_0 = v_0^3, \varepsilon = 10^{-4}$, $J = 159.0, u_{\max} = 13.63$. So far, the existence of such a positive solution is still an open problem.*
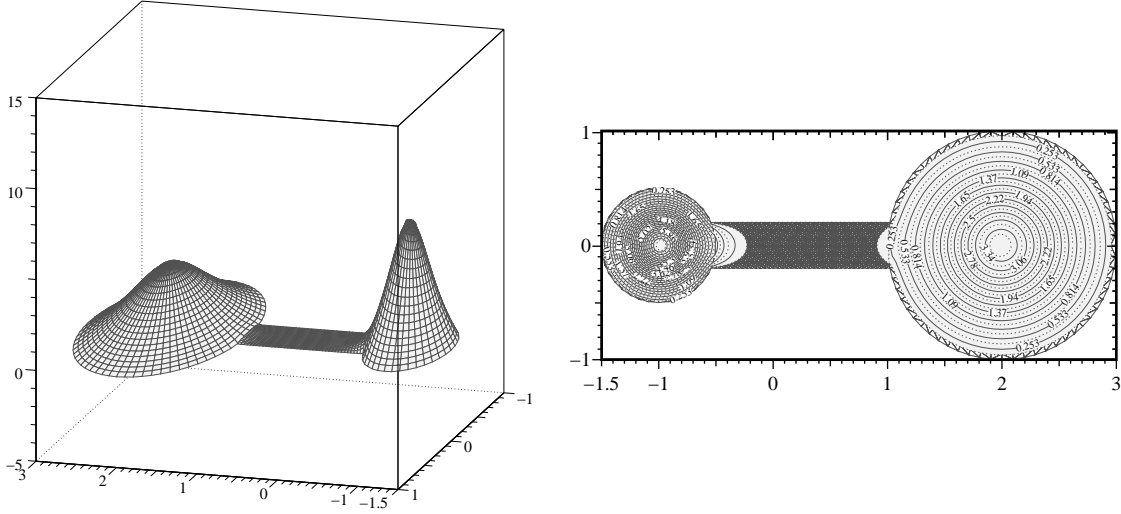
FIG. 5. *A solution w_2 with MI $= 2$ and its contours.* $L = [w_1^1], v_0 = v_0^2$, $\varepsilon = 10^{-4} \times 6, J = 53.12, u_{max} = 7.037$.

If we use $L = [w_1^1, w_2]$ and $v_0 = v_0^3$ to search for a solution with MI $= 3$, the algorithm yields a solution with positive and negative peaks. This can be explained as follows. Since a function with a larger energy value becomes less stable and a solution with a larger MI is also less stable. Note that

$$J(w_1^1) < J(w_1^2) < J(w_1^3).$$

When we use $L = [w_1^1, w_2]$ and $v_0 = v_0^3$ to search for a solution with MI $= 3$, we start the process at searching for a peak with lower energy for a solution with a lower MI and then go to search for a peak with larger energy for a solution with a higher MI. The process becomes very unstable. Now if we switch the order, we start the process at searching for a peak with higher energy for a solution with a lower MI and then go to next stage to search for a peak with lower energy for a solution with a higher MI. The stability of the process is balanced. Thus when we use $L = [w_1^3]$ and $v_0 = v_0^2$ to find a solution w_2^2 with MI $= 2$ and two positive peaks in the left compartment and the central corridor. Then we use $L = [w_1^3, w_2^2]$ and $v_0 = v_0^1$ to search for a positive solution with MI $= 3$, we obtain
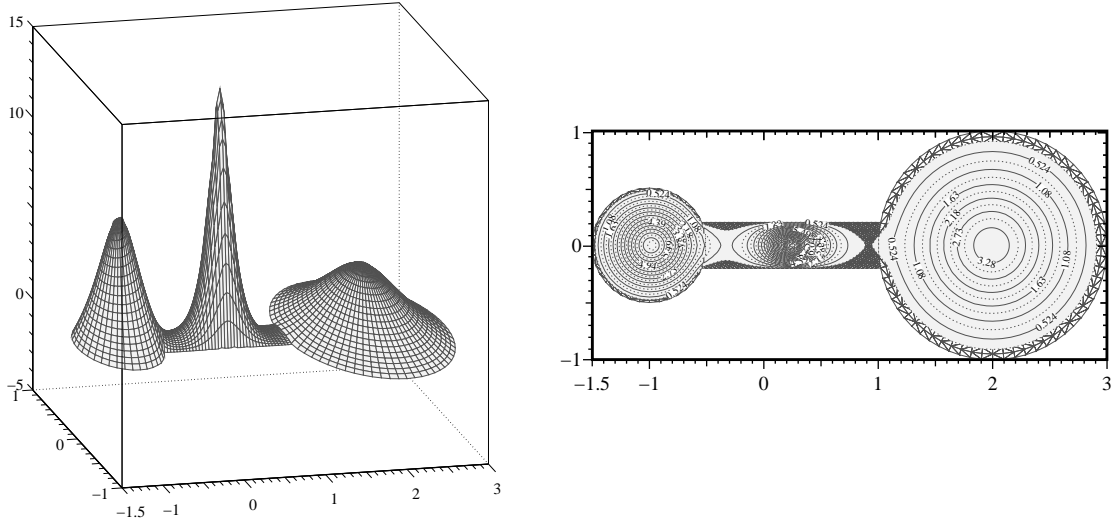
FIG. 6. *A solution w_3 with MI = 3 and its contours. $\varepsilon = 10^{-3}, J = 212.5$, $u_{\max} = 13.78$. This is the only positive solution with MI = 3 we can find.*

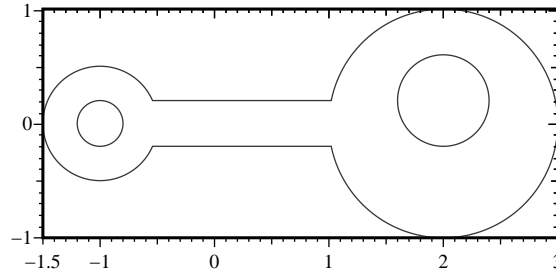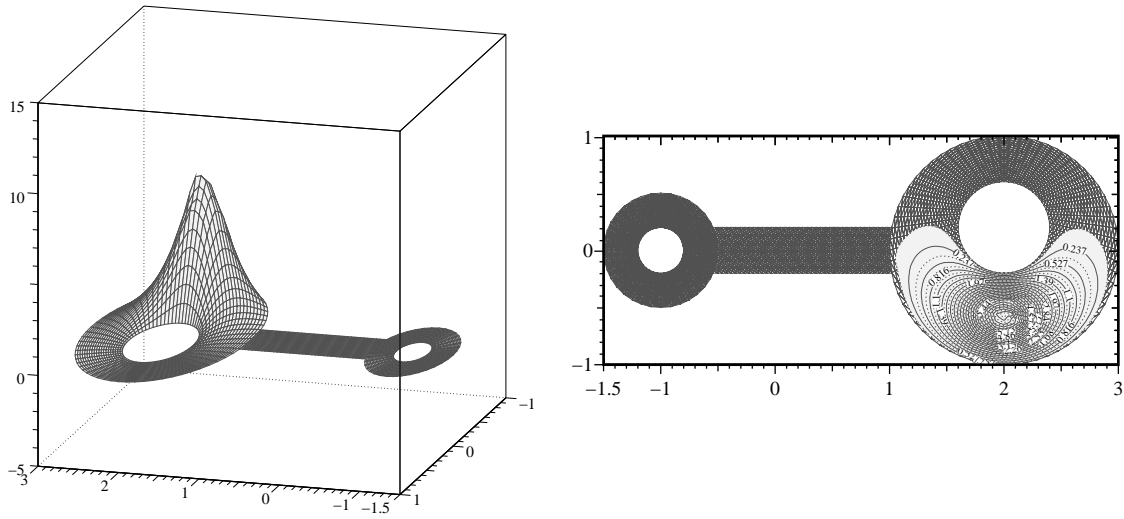**Case 2:** On a dumbbell-shaped domain with two cavities.

FIG. 7. *A dumbbell-shaped domain with two cavities.*

We use, respectively, the following four "mound" functions as initial ascent directions.

$$v_0^i(x) = \begin{cases} \cos(\dfrac{|x - x_i|}{d_i}\dfrac{\pi}{2}) & \text{if } |x - x_i| \leq d_i, \\ 0 & \text{otherwise,} \end{cases}$$

with

$$x_1 = (2, -0.6), \quad d_1 = 0.4, \quad x_2 = (-0.5, 0), \quad d_2 = 0.2,$$
$$x_3 = (0.25, 0), \quad d_3 = 0.2, \quad x_4 = (-1.35, 0), \quad d_4 = 0.15.$$

FIG. 8.   *The ground state solution w_1^1 with $MI = 1$ and its contours.* $v_0 = v_0^1$, $\varepsilon = 10^{-4}$, $J = 44.18$, $u_{\max} = 6.664$.
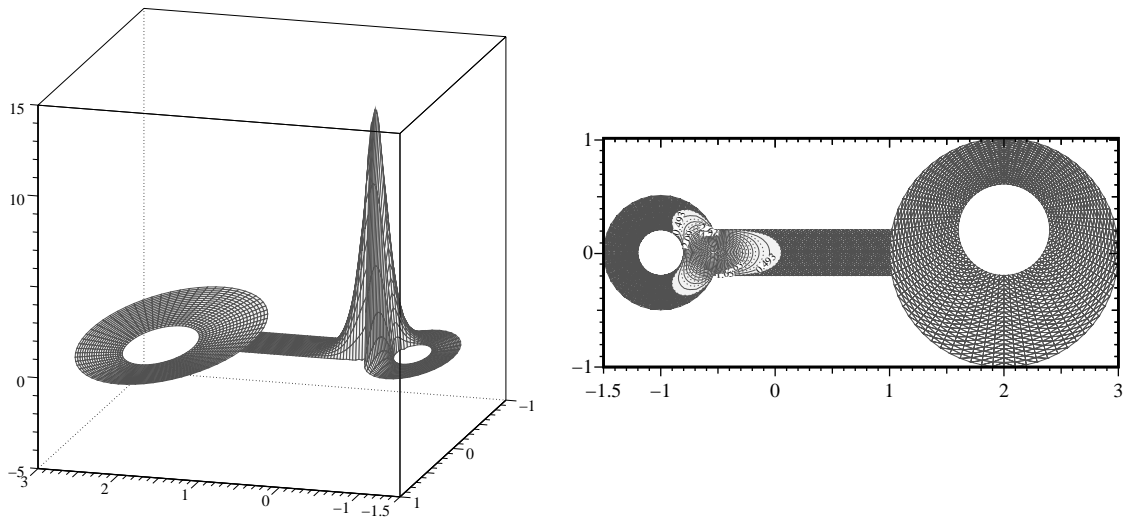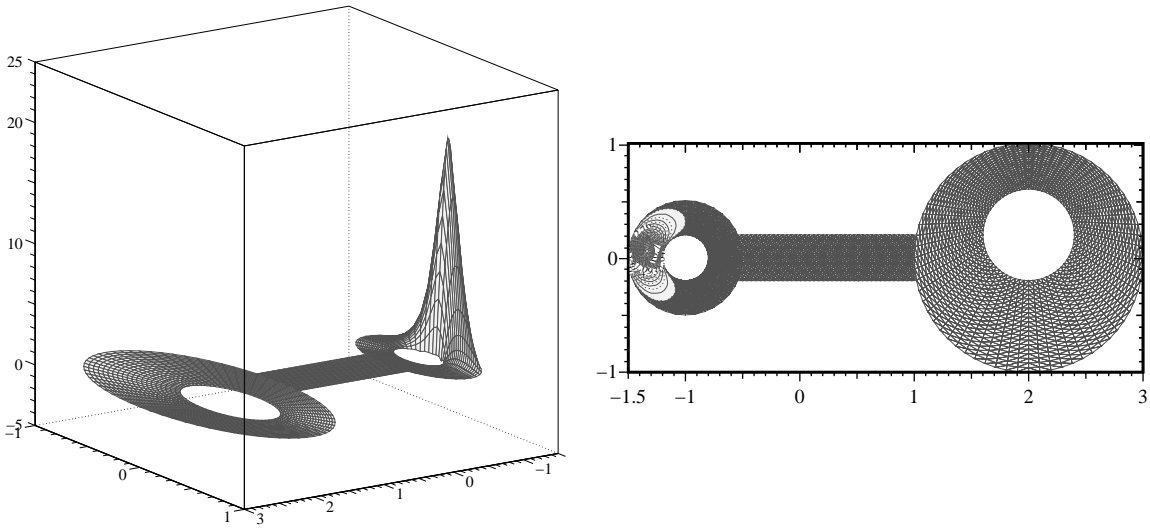

FIG. 9.   *The second solution w_1^2 with $MI = 1$ and its contours.* $v_0 = v_0^3$, $\varepsilon = 10^{-4}$, $J = 153.5$, $u_{\max} = 12.94$.
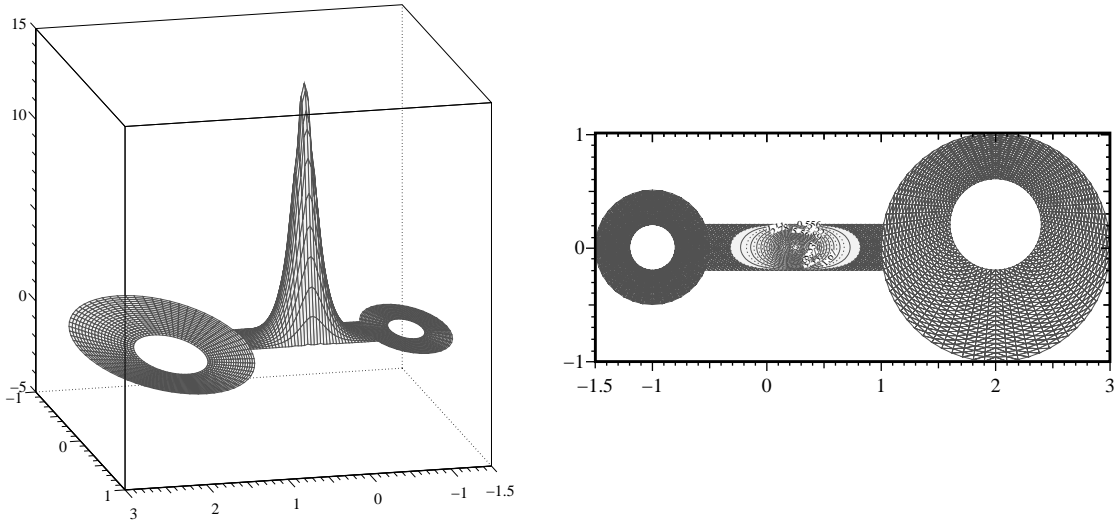
FIG. 10. The third solution w_1^3 with $MI = 1$ and its contours. $v_0 = v_0^2$, $\varepsilon = 10^{-4}, J = 165.6, u_{max} = 13.64$. Its profile is similar to w_1^3 in Figure 4. So far, the existence of such a positive solution is still an open problem.

FIG. 11. The fourth solution w_1^4 with $MI = 1$ and its contours. $v_0 = v_0^4$, $\varepsilon = 10^{-4}, J = 286.1, u_{max} = 17.85$.
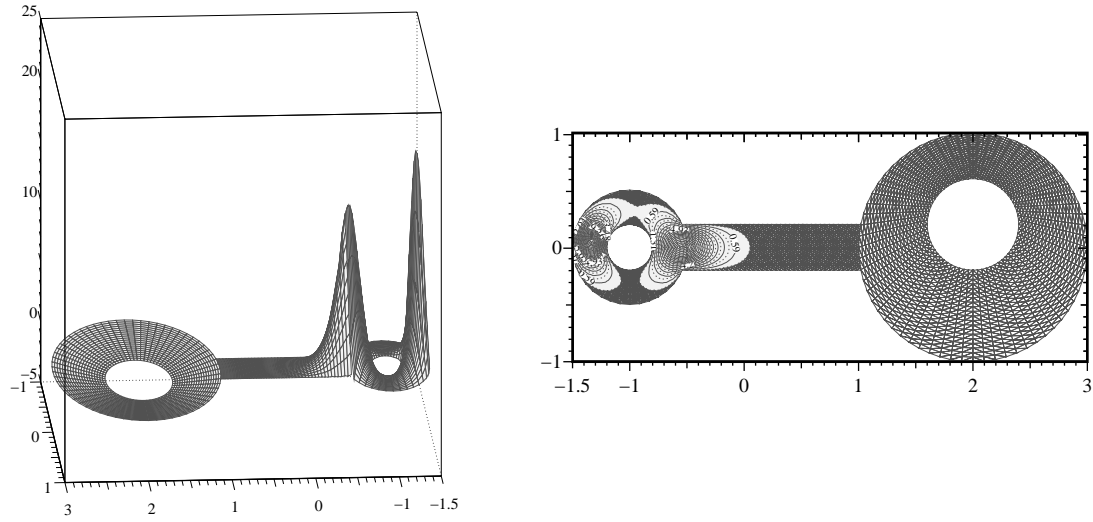
FIG. 12.   A solution w_2 with $MI = 2$ and its contours. $L = [w_1^4], v_0 = v_0^2,$ $\varepsilon = 10^{-4} \times 6, J = 439.5, u_{\max} = 17.85$. There are other positive solutions with $MI = 2$.
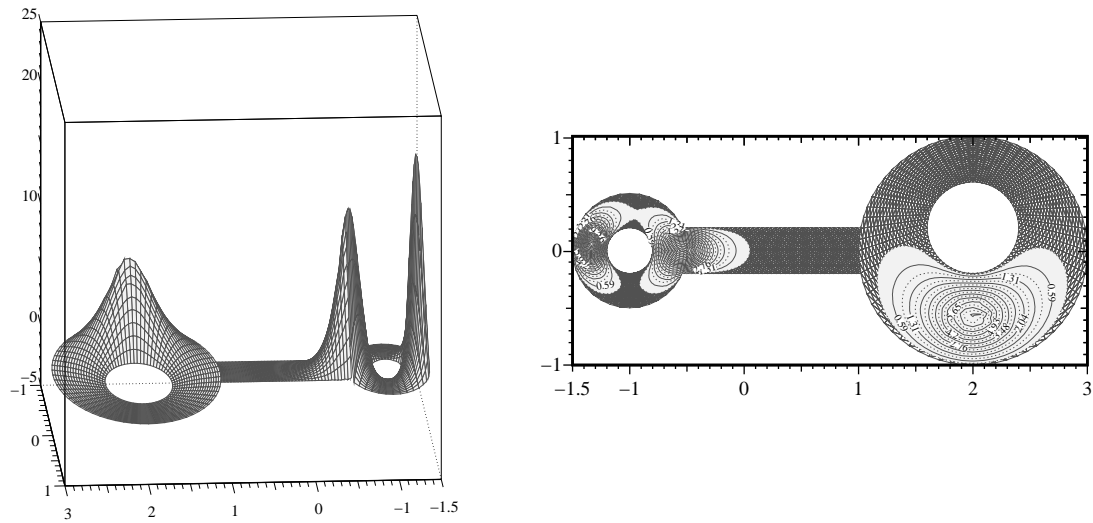

FIG.   13.   A solution w_3 with $MI = 3$ and its contours. $L = [w_1^4, w_2], v_0 = v_0^1,$ $\varepsilon = 10^{-3}, J = 483.6, u_{\max} = 17.86$. There are other positive solutions with $MI = 3$.

Case 3: On a concentric annulus with inner radius = 0.7 and outer radius = 1.

The domain is a nice geometric figure. However, due to the symmetry, any solution being rotated for any angle is still a solution. Thus each solution belongs to an one parameter family of solutions. For this case, the existence of non radially symmetric positive solutions has been established by Coffman [13] and Li [22]. The number of positive peaks that a solution may have depends on the width of the annulus. If we utilize the symmetry, we can find a radially symmetric solution that turns out to be a local (global) maximum. Otherwise, this case is highly degenerate. When the boundary is discretized into a polygon, theoretically the case becomes nondegenerate. However, when the discretization is fine, each solution has other solutions nearby. The computation becomes even tougher. After several iterations, there are multiple solutions inside a small neighborhood of the numerical solution. The algorithm may start to wander around. We have used 384 elements on the outer circle, 192 elements on the inner circle and the following "mound" function as an initial ascent direction for $\theta_i = \frac{(i-1)\pi}{4}$,

$$v_0^i(x) = \begin{cases} \cos(\dfrac{|x - 0.85(\cos\theta_i, \sin\theta_i)|}{0.15} \dfrac{\pi}{2}) & \text{if } |x - 0.85(\cos\theta_i, \sin\theta_i)| \leq 0.15, \\ 0 & \text{otherwise.} \end{cases}$$
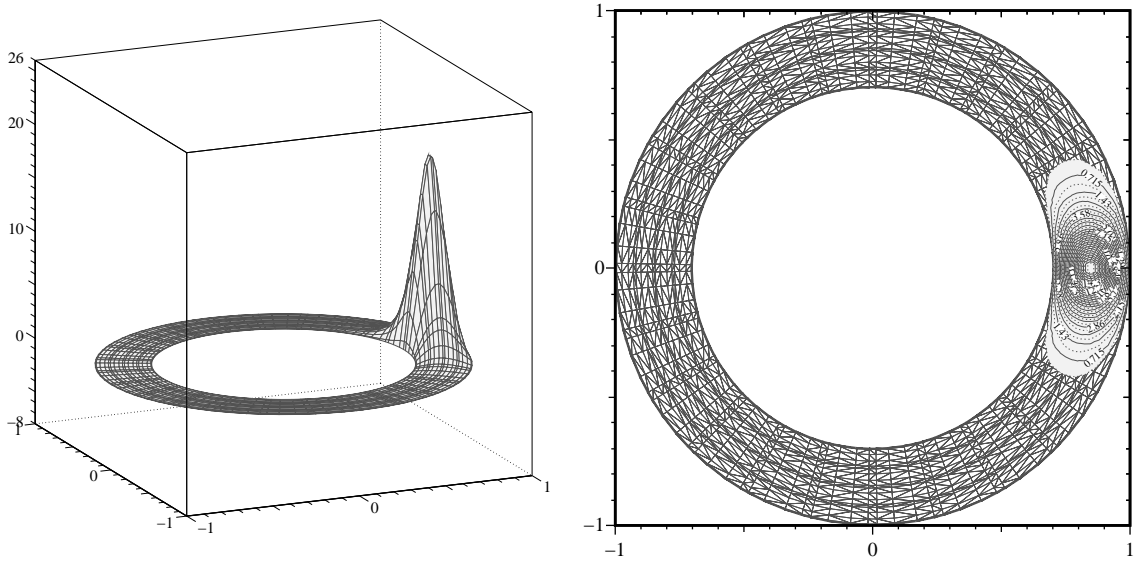

FIG. 14. *A ground state solution* w_1 *with MI = 1 and its contours.* $v_0 = v_0^1$, $\varepsilon = 10^{-4}, J = 289.1, u_{\max} = 18.12$.
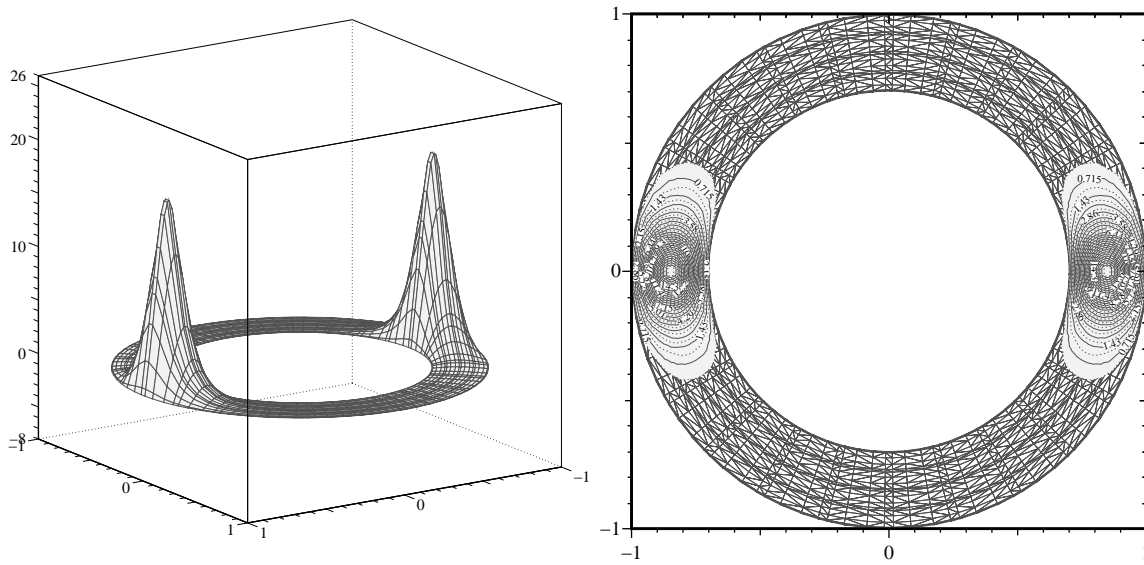
FIG. 15.   *A solution w_2 with $MI = 2$ and its contours.* $L = [w_1], v_0 = v_0^3,$
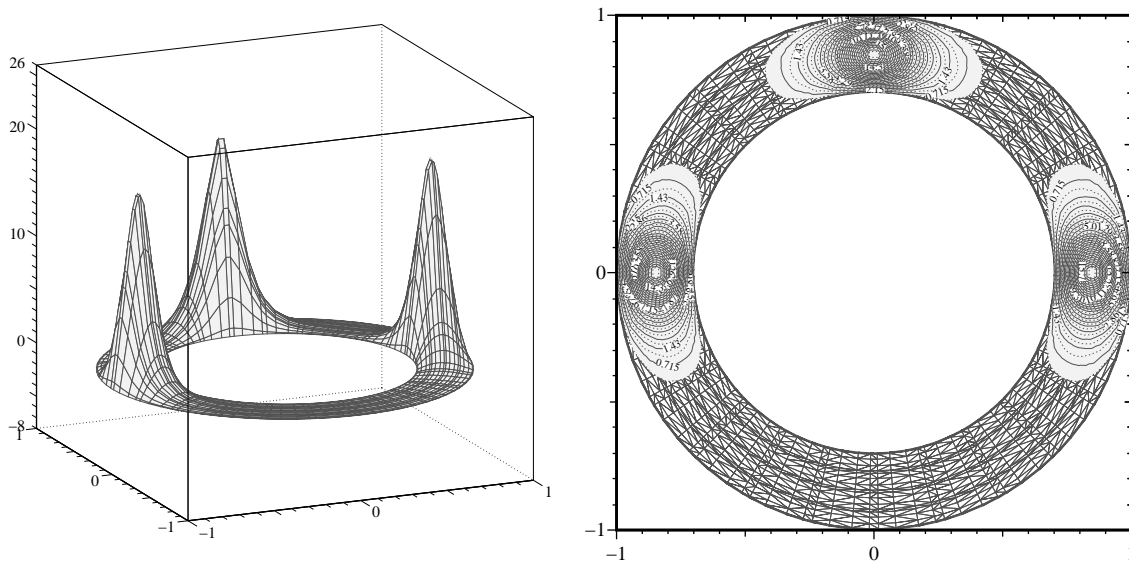$\varepsilon = 10^{-4}, J = 579.0, u_{\max} = 18.12.$

FIG. 16.   *A solution w_3 with $MI = 3$ and its contours.* $L = [w_1, w_2], v_0 = v_0^2,$
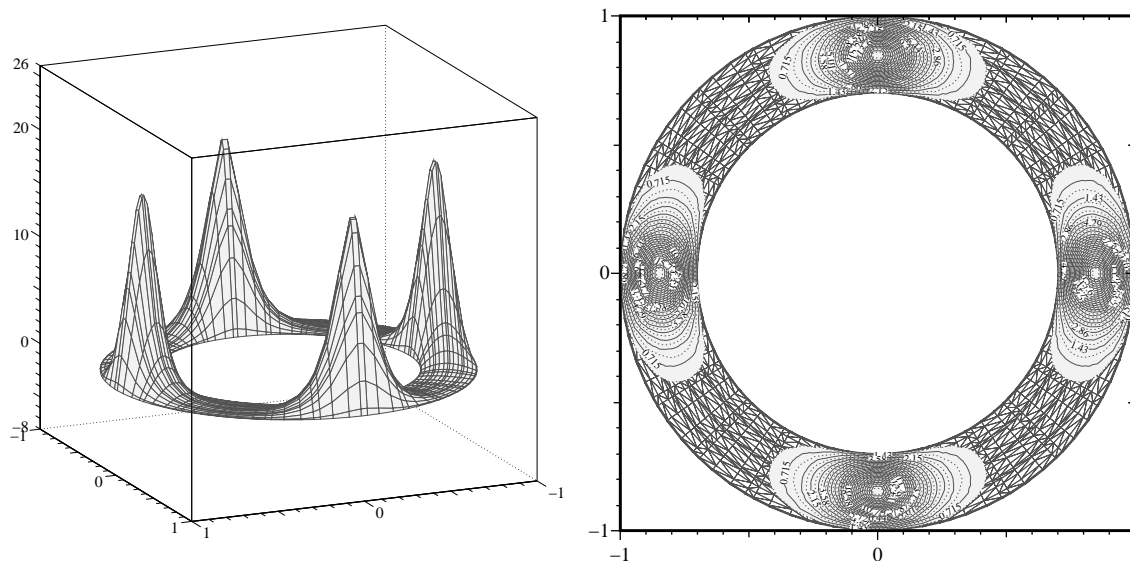$\varepsilon = 10^{-4} \times 6, J = 868.8, u_{\max} = 18.12.$

Fig. 17. *A solution w_4 with $MI = 4$ and its contours. $L = [w_1, w_2, w_3]$, $v_0 = v_0^4$, $\varepsilon = 10^{-3} \times 3$, $J = 1159$, $u_{\max} = 18.12$.*

Finding multiple saddle points is important for both theory and applications. However it is very challenging. Little is known in the literature. We try to develop some numerical algorithms and corresponding mathematical theory for finding such saddle points in a stable way. It is known that many saddle points can not be approximated. One can only numerically approximate those multiple saddle points with some "nice" properties, e.g., minimax solutions. We classify those saddle points through mathematical analysis. The results presented in this paper are under some "reasonably nice" conditions. They provide a mathematical foundation for our further research. Meanwhile those conditions will be further generalized. Methods to check those conditions (e.g., the continuity or differentiability condition of p) will be developed as research in this direction progresses.[1] So far, the algorithm is still better than mathematical analysis. It produced many interesting numerical results that beyond theoretical results. For example, when the profiles of solutions in Figures 4 and 10 are presented to nonlinear PDE analysts in 1997 and 1998, they generated warm debates about the existence and the Morse indices of such solutions. We are pleased to know that some results on the existence of such solutions has been recently proved (See [35]). Our algorithm can be used to solve for a critical point which is not a minimax solution, e.g., a Monkey saddle point. However the analysis is beyond the scope of any minimax principle, more profound approach is required. As mathematical analysis in this research progresses, the algorithm will be accordingly modified. Convergence is a paramount issue of any numerical algorithm. The Morse index of a solution is an important notion that provides understanding of the local structure of a saddle points and can be used to measure instability of a saddle point. Although, in the above numerical examples, we have printed the Morse index for each

---

[1]Results in this project have been obtained and will be presented in a future paper.

numerical solution, it is based on the way we compute the solution in the algorithm, its mathematical verification has not been established. Due to the limitation to the length of this paper, results on those issues will be addressed in a subsequent paper [20] and future papers [21, 37].

References

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.

[2] H. Amann, Supersolution, monotone iteration and stability, *J. Diff. Eq.* 21 (1976), 367–377.

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

[4] J. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley, New York, 1984

[5] T. Bartsch and Z.Q. Wang, On the existence of sign-changing solutions for semilinear Dirichlet problems, *Topol. Methods Nonlinear Anal.* 7 (1996), 115–131.

[6] V. Benci and G. Cerami, The effect of the domain topology on the number of positive solutions of nonlinear elliptic problems, *Arch. Rational Mech. Anal.* 114(1991), 79–94.

[7] L. Bieberbach, $\Delta u = e^u$ und die automorphen Functionen, *Mathematische Annalen* 77(1916), 173-212.

[8] H. Brezis and L. Nirenberg, Remarks on Finding Critical Points, *Communications on Pure and Applied Mathematics*, Vol. XLIV, 939-963, 1991.

[9] A. Castro, J. Cossio and J. M. Neuberger, A sign-changing solution for a superlinear Dirichlet problem, *Rocky Mountain J. Math.* 27(1997), 1041-1053.

[10] K.C. Chang, *Infinite Dimensional Morse Theory and Multiple Solution Problems*, Birkhäuser, Boston, 1993.

[11] G. Chen, W. Ni and J. Zhou, Algorithms and Visualization for Solutions of Nonlinear Elliptic Equations Part I: Dirichlet Problems, *Int. J. Bifurcation & Chaos*, to appear.

[12] 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.

[13] C.V. Coffman, A nonlinear boundary value problem with many positive solutions, *J. Diff. Eq.* 54(1984), 429-437.

[14] E.N. Dancer, The effect of domain shape on the number of positive solutions of certain nonlinear equations, *J. Diff. Eq.* 74 (1988), 120–156.

[15] Y. Deng, G. Chen, W.M. Ni, and J. Zhou, Boundary element monotone iteration scheme for semilinear elliptic partial differential equations, *Math. Comp.* 65 (1996), 943–982.

[16] 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),

[17] 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.

[18] D. Greenspan and S.V. Parter, Mildly nonlinear elliptic partial differential equations and their numerical solution, II, *Numer. Math.* 7 (1965), 129–146.

[19] K. Ishihara, Monotone explicit iterations of the finite element approximations for the nonlinear boundary value problems, *Numer. Math.* 45 (1984), 419–437.

[20] Y. Li and J. Zhou, Convergence results of a minimax method for finding critical points, in review.

[21] Y. Li and J. Zhou, Local characterizations of saddle points and their Morse indices, *Advances in Control of Nonlinear Distributed Parameter Systems* , Marcel Dekker, New York, pp. 233-252, to appear.

[22] Y.Y. Li, Existence of many positive solutions of semilinear elliptic equations on annulus, *J. Diff. Eq.* 83 (1990), 348–367.

[23] F. Lin and T. Lin, Minimax solutions of the Ginzburg-Landau equations, *Slecta Math. (N.S.)*, 3(1997) no. 1, 99-113.

[24] J. Mawhin and M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer-Verlag, New York, 1989.

[25] Z. Nehari, On a class of nonlinear second-order differential equations, *Trans. Amer. Math. Soc. 95* (1960), 101-123.

[26] W.M. Ni, *Some Aspects of Semilinear Elliptic Equations*, Dept. of Math. National Tsing Hua Univ., Hsinchu, Taiwan, Rep. of China, 1987.

[27] W.M. Ni, Recent progress in semilinear elliptic equations, in RIMS Kokyuroku 679, Kyoto University, Kyoto, Japan, 1989, 1-39.

[28] C.V. Pao, *Nonlinear Parabolic and Elliptic Equations*, Plenum Press, New York, 1992.

[29] S.V. Parter, Mildly nonlinear elliptic partial differential equations and their numerical solutions I, *Numer. Math.* 7 (1965), 113–128.

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

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

[32] S. Shi, Ekeland's Variational Principle and the Mountain Pass Lemma, *Acta Mathematica Sinica,* Vol. 1 No. 4, 348-355, 1985.

[33] M. Struwe, *Variational Methods*, Springer, 1996.

[34] Z. Wang, On a superlinear elliptic equation, *Ann. Inst. Henri Poincare*, 8(1991), 43-57.

[35] J. Wei and L. Zhang, "On the effect of the domain shape on the existence of large solutions of some superlinear problems", preprint.

[36] M. Willem, *Minimax Theorems*, Birkhauser, Boston, 1996.

[37] J. Zhou, Instability indices of saddle points by a local minimax method, preprint.