Optimization with Some Uncontrollable Variables: A Min-Equilibrium Approach

Jianxin Zhou*

Texas A&M University, College Station, Texas, USA

Keywords: Minimization with some uncontrollable variables, riskiness index, saddle critical point analysis and computation.

Abstract

Motivated by instability analysis of unstable (excited state) solutions in computational physics/chemistry, in this paper, the minimax method for solving an optimal control problem with partially uncontrollable variables is embedded into a more general min-equilibrium problem. Results in saddle critical point analysis and computation are modified to provide more information on the minimized objective values and their corresponding riskiness for one to choose in decision making. A numerical algorithm to compute such minimized objective values and their corresponding riskiness is devised. Some convergence results of the algorithm are also established.

1 Introduction

In the study of self-guided light waves in nonlinear optics [1,2,6], solutions that are not ground states are the excited states and called solitons, are of great interests. All solitons are saddle critical point type unstable solutions. Among various modes of solitons are the vortex-mode and dipole-mode vector solitons. It has been experimentally and numerically observed

*Supported in part by NSF Grant DMS-0311905. E-mail address: jzhou@math.tamu.edu

that different solitons have different responses to perturbation, e.g., the vortex-mode can be easily perturbed to decay into a dipole-mode; While the dipole-modes are much more stable, "stable enough for experimental observation, ..., extremely robust, have a typical lifetime of several hundred diffraction lengths and survive a wide range of perturbations"[1].

Motivated by the above observation, let us consider the following optimal control problem:

(1.1) One wishes to locally minimize an objective functional $J : X \times Y \to \mathbb{R}$ where X represents the space of controllable variable x and Y represents the space of variable y uncontrollable upto small perturbations.

For this class of problems, to get a reliable solution, *the most conservative strategy* widely used in the optimal control literature is to seek a minimax solution of the form

$$u^* = (x^*, y^*) = \arg \min_{x \in X} \max_{y \in Y} J(x, y).$$

In this paper, since we only consider local minimization and small perturbations, all min and max are in the local sense. It is quite natural to ask the following questions:
(a) Why one has to be most conservation? If there is another solution with a much larger gain and a small riskiness, one may be willing to pursue such a solution.
(b) What do we mean by other *solutions* and how to find them?

In this paper, we shall answer the questions by using critical point theory and method to embed the original problem (1.1) into a min-equilibrium problem with multiple solutions, each of which has an objective value and a riskiness index attached. Thus one can compare those solutions' objective values and riskiness indices, and select the best solution in a decision making process.

First let us discuss what kind of "solutions" we wish to find. Let (x^*, y^*) be a "solution". Since the y-variable is uncontrollable upto a small perturbation Δy, we hope that Δy will not dramatically change the value of $J(x^*, y^*)$. Thus we assume $J_y(x^*, y^*) = 0$. Then J is minimized over the x-variable at x^* implies $J_x(x^*, y^*) = 0$, i.e., a "solution" (x^*, y^*) should satisfy $\nabla J(x^*, y^*) = (J_x(x^*, y^*), J_y(x^*, y^*)) = 0$.

We need some related notions from critical point theory. Let $H = X \times Y$ be a Hilbert space and $J : H \to \mathbb{R}$ be a C^1-function. A point $u^* = (x^*, y^*) \in H$ is a *critical point* of J, if

u^* solves its Euler-Lagrange equation

$$\nabla J(u) = (J'_x(x,y), J'_y(x,y)) = 0, \tag{1.2}$$

where $\nabla J(u)$ is the Frechet gradient of J at $u \in H$. In physical systems, critical points appear as *equilibrium states*. Thus any equilibrium state $u = (x, y)$ of J which minimizes J over the x-variable is called a *min-equilibrium solution* or a *solution* to the problem (1.1).

It is clear that if $u^* = (x^*, y^*)$ is a minimax solution, then u^* is also a min-equilibrium solution and in this case, a small perturbation in the y-variable will only decrease the value of J, which is certainly welcome. Thus it is the most reliable solution or takes no risk. But its objective value is also expected to be the largest among other solutions. So it is the most conservative solution. While other min-equilibrium solutions do not maximize J over the y-variable, which is uncontrollable upto a small perturbation Δy. We may expect that those solutions have smaller objective values but risk to have a possibility to increase their objective values due to Δy, and lost reliability on their optimality.

It is interesting to ask: can we measure such riskiness (reciprocal to the reliability) involved in each of the other solutions? If the answer is yes, then in a decision making process, one can compare the values of J at those solutions with their riskiness, and select the best solution by balancing its value and riskiness.

In Section 2, we establish a min-equilibrium approach. Riskiness analysis will be done in Section 3. A numerical algorithm for solving the problem and its convergence analysis will be presented in Section 4. Section 5 will be used to discuss some related theoretical issues.

2 A Local Min-Equilibrium Method

The new approach, called a min-equilibrium approach, is to embed the original minimax method into a more general approach of finding equilibrium states of J where the values of J are minimized over the x-variable. This approach unifies two minimax methods in optimal control theory and in critical point theory. Thus results in saddle critical point analysis and computation can be applied.

DEFINITION 2.1. *A set-valued mapping $Q : X \to 2^Y$ is called the* y-equilibrium mapping *of J if $Q(x) = \{y \in Y : J'_y(x,y) = 0\}$. A single-valued mapping $q : X \to Y$ is a* y-equilibrium selection *of J if $q(x) \in Q(x), \forall x \in X$. If q is locally defined, then q is called a* local y-equilibrium selection *of J.*

Then it is direct to verify that

LEMMA 2.1. *If J is C^1, then the graph $\mathbb{G} = \{(x,y) : x \in X, y \in Q(x) \neq \emptyset\} = \{(x,y) : J'_y(x,y) = 0\}$ is closed.* ▮

THEOREM 2.1. *Let q be a local y-equilibrium selection of J at $x_0 \in X$ s.t. q is continuous at x_0. If $J'_x(x_0, q(x_0)) \neq 0$, then there exists $s_0 > 0$ s.t. when $0 < s < s_0$,*

$$J(x(s), q(x(s))) < J(x_0, q(x_0)) - \frac{s}{2}\|J'_x(x_0, q(x_0))\|^2 \tag{2.1}$$

where $x(s) = x_0 - sJ'_x(x_0, q(x_0))$.

Proof. Since $x(s) \to x_0$ and $q(x(s)) \to q(x_0)$ as $s \to 0$ and J is C^1, by denoting $z(s) = (x(s), q(x(s))) - (x_0, q(x_0))$, we have

$$\begin{aligned} & J(x(s), q(x(s))) - J(x_0, q(x_0)) = \langle \nabla J(x_0, q(x_0)), z(s) \rangle_H + o(\|z(s)\|) \\ = & \langle J'_x(x_0, q(x_0)), x(s) - x_0 \rangle_X + o(\|z(s)\|) = -s\|J'_x(x_0, q(x_0))\|^2 + o(\|z(s)\|) \\ < & -\frac{s}{2}\|J'_x(x_0, q(x_0))\|^2, \quad \text{as } 0 < s < s_0, \end{aligned}$$

where the last inequality holds because $z(s) \to 0$ as $s \to 0$ and $-s\|J'_x(x_0, q(x_0))\|^2 < 0$. ▮

Inequality (2.1) will be used to define a stepsize rule in a numerical algorithm presented in Section 4. As a direct consequence of Theorem 2.1, we obtain the following characterization of a local min-equilibrium solution.

THEOREM 2.2. *Let q be a local y-equilibrium selection of J at $x^* \in X$ s.t. q is continuous at x^*. If $x^* = \arg\min_{x \in X} J(x, q(x))$, then $(x^*, q(x^*))$ is a min-equilibrium solution to (1.1).* ▮

Proof. By the definition of q, $J_y(x^*, q(x^*)) = 0$. Since $x^* = \arg\min_{x \in X} J(x, q(x))$, Theorem 2.1 implies that $J_x(x^*, q(x^*)) = 0$. Thus $\nabla J(x^*, q(x^*)) = 0$ and J is already minimized over the x-variable. Therefore $(x^*, q(x^*))$ is a min-equilibrium solution to (1.1). ▮

3 Riskiness Analysis of Min-Equilibrium Solutions

Since the y-variable in J is uncontrollable upto a small perturbation Δy, a min-equilibrium state $u^* = (x^*, y^*)$ risks to have a possibility to increase the value of J at (x^*, y^*) due to Δy. Thus it is interesting to study how to measure such riskiness.

For each $x \in X$, let $G_x(y) \equiv J(x,y)$. Then $G_x : Y \to \mathbb{R}$. Let $y^* \in Y$ be a critical point of G_x, i.e., $G'_x(y^*) = 0$. The first candidates for critical points are the local extrema which the classical calculus of variation devoted to and traditional numerical algorithms focused on. Critical points that are not local extrema are called *saddle points*. In physical systems, saddle points appear as *unstable equilibrium* or excited states, whose instability can be defined in several different ways.

A vector $y \in Y$ is an *increasing direction* of G_x at y^* if there is $T > 0$ s.t.

$$G_x(y^* + ty) > G_x(y^*) \quad \forall\, T > |t| > 0.$$

The dimension of a maximum subspace of such increasing directions of G_x at y^* is called the *local saddle index* (LSI) of G_x at y^*. Since such an index lacks of characterization and is too difficult to compute. Let us study other alternatives.

The Morse Index. If $G''_x(y^*)$ is a self-adjoint Fredholm operator, then the spectral decomposition $\; Y = Y^- \oplus Y^0 \oplus Y^+ \;$ exists where Y^-, Y^+ and Y^0 are respectively the maximum negative, positive definite, the null subspaces of $G''_x(y^*)$ in Y with $\dim(Y^0) < \infty$. The integer $\dim(Y^+)$ is called the *Morse index* (MI) of y^*. If y^* is a nondegenerate critical point of G_x, i.e., $Y^0 = \{0\}$, then it is clear that

$$\text{MI}\,(y^*) = \text{LSI}\,(y^*).$$

The Morse index is commonly used in the literature to measure the instability of unstable solutions [8, 13]. But MI is very expensive to compute and not useful in handling degenerate cases due to the fact that many different situations may happen in Y^0 and the Morse index is inclusive about Y^0.

Borrowing a notion called *the order of saddle* from computational chemistry/physics [3], we introduce the following definition, also called *local saddle index* and use it to define *the local riskiness index* (LRI).

Let $y^* \in Y$. If $Y = Y^I \oplus Y^D$ for some subspaces Y^I, Y^D in Y s.t. for each $y_1 \in Y^I$ and $y_2 \in Y^D$ with $\|y_1\| = \|y_2\| = 1$, there exist $r_1 > 0$ and $r_2 > 0$ satisfying

$$G_x(y^* + ty_1 + o(t)) \;>\; G_x(y^*),\ \forall\, 0 < |t| \le r_1, \tag{3.1}$$

$$G_x(y^* + ty_2) \;<\; G_x(y^*),\ \forall\, 0 < |t| \le r_2, \tag{3.2}$$

where $o(t)$ represents a higher order term. Then y^* is a saddle point of G_x and $\dim(Y^I)$ is called the *local saddle index* (LSI) of y^*.

Thus for G_x, a critical point of LSI=0 is a local maximum point. If y^* is nondegenerate and $|t| > 0$ is small, then we have

$$\begin{aligned} G_x(y^* + ty_1 + o(t)) - G_x(y^*) &= \frac{t^2}{2}\langle G''_x(y^*)y_1, y_1\rangle + o(t^2) > 0, \quad \forall y_1 \in Y^+, \\ G_x(y^* + ty_2) - G_x(y^*) &= \frac{t^2}{2}\langle G''_x(y^*)y_2, y_2\rangle + o(t^2) < 0, \quad \forall y_2 \in Y^-, \end{aligned}$$

i.e., $Y^D = Y^-$, $Y^I = Y^+$, or $\mathrm{LSI}\,(y^*) = \mathrm{MI}\,(y^*)$. But (3.1) and (3.2) do not relate to degeneracy. Thus LSI is more general than MI.

Definition 3.1. *If $u^* = (x^*, y^*)$ is a min-equilibrium solution of J, i.e.,*

$$u^* = (x^*, y^*) = \arg\min_{x\in X} J(x, q(x))$$

and y^* is a saddle point of $G_{x^*}(\cdot) \equiv J(x^*, \cdot)$ with LSI(y^*), then the min-equilibrium solution u^* of J is said to have a local riskiness index *(LRI) equal to LSI(y^*).*

Thus to find LRI of a min-equilibrium state (x^*, y^*) of J is equivalent to computing LSI of G_{x^*} at y^*. So we may just focus on computing LSI of G_x at y^*. Since the case LRI=0 corresponding to the minimax solution is clear, we study the case where LRI≥ 1.

To adopt the approach developed in [4, 5, 13, 14], for each $x \in X$, we use a translation and denote $G_x(y) \equiv J(x, y^*_x + y)$ where y^*_x is a local maximum point of $J(x, \cdot)$. Thus G_x has a local maximum at 0. Let L be a closed subspace of Y and denote $S_{L^\perp} = \{y \in Y : \|y\| = 1, y \perp L\}$. For each $y \in S_{L^\perp}$, denote $[L, y] = \{ty + y_L : t \in \mathbb{R}, y_L \in L\}$.

Definition 3.2. *([4, 5]) A set-valued mapping $P\colon S_{L^\perp} \to 2^Y$ is called the* trough *mapping of G_x if $P(y)$ contains all local minimum points of G_x in $[L, y]$. A mapping $p : S_{L^\perp} \to Y$ is called a* trough *selection of G_x if $p(y) \in P(y)$ for each $y \in S_{L^\perp}$. If p is locally defined, then p is called a* local trough *selection of G_x.*

Note that the above definition depends on the x-variable and by [4, 5], a max-min solution

$$y^* = \arg\max_{y\in S_{L^\perp}} G_x(p(y)) = \arg\max_{y\in S_{L^\perp}} \min_{z\in[L,y]} G_x(z)$$

yields a saddle critical point $z^* = p(y^*)$ of G_x and furthermore

THEOREM 3.1. *([13]) If p is a local trough selection of G_x differentiable at $y^* \in S_{L^\perp}$, $p(y^*) \notin L$ and $y^* = \arg\max_{y\in S_{L^\perp}} G_x(p(y))$, then $z^* = p(y^*)$ is a saddle point G_x with*

$$\dim(L) + 1 = MI(z^*) + \dim(Y^0 \cap \{L, y^*\}). \quad \blacksquare \tag{3.3}$$

THEOREM 3.2. *([14]) Let p be a local trough selection of G_x differentiable at $y^* \in S_{L^\perp}$, $p(y^*) \notin L$ and $y^* = \arg\max_{y\in S_{L^\perp}} G_x(p(v))$. If the local max with p is strict, then $z^* \equiv p(v^*)$ is a saddle point of G_x with $LSI = \dim(L) + 1$.* $\blacksquare$

It is clear that the subspace L in the above discussion depends on the x-variable as well. Thus when we consider a fixed riskiness level k, we use L_x s.t. $\dim(L_x) = k$ for all $x \in X$.

Consequently, for the optimal control problem under consideration

(1) let

$$(x^*, y^*) = \arg\min_{x\in X}\max_{y\in Y} J(x,y),$$

then (x^*, y^*) is a solution with LRI $(x^*, y^*) = 0$;

(2) let $L_x = [0]$ and

$$(x^*, y^*) = \arg\min_{x\in X}\max_{y\in S_{L^\perp}} G_x(p(y)) = \arg\min_{x\in X}\max_{y\in S_{L^\perp}}\min_{z\in[y]} J(x, y^*_x + z),$$

then (x^*, y^*) is a solution with LRI $(x^*, y^*) = 1$;

(3) in general, for given $k = 1, 2, ...$, let $L_x = [0, y^1_x, ..., y^{k-1}_x]$ where y^i_x, $i = 1, ..., k-1$ are saddle points of $G_x(\cdot) \equiv J(x, y^*_x + \cdot)$ with LSI $(y^i_x) = i$ and computed from previously found saddle points $y^1_x, ..., y^{i-1}_x$ with LSI$= 1, ..., i-1$. If

$$(x^*, y^*) = \arg\min_{x\in X}\max_{y\in S_{L_x^\perp}} G_x(p(y)) = \arg\min_{x\in X}\max_{y\in S_{L_x^\perp}}\min_{z\in[L_x,y]} J(x, y^*_x + z),$$

then (x^*, y^*) is a solution with LRI $(x^*, y^*) = \dim(L_x) + 1 = k$.

Let $k = 0, 1, ...$ be given and fixed. For each $x \in X$, $p(y)$ and L_x as given in the above (2)-(3), if we let

$$q(x) = \{y^*_x + z_x | z_x = \arg\max_{y\in S_{L_x^\perp}}\min_{z\in[L_x,y]} J(x, y^*_x + z)\},$$

then it is clear that $(J'_x(x, q(x)), J'_y(x, q(x))) = 0$, i.e., such $q(\cdot)$ defines a y-equilibrium selection of J. Hence Theorems 2.1 and 2.2 can be applied. Let

$$\mathcal{M} = \{(x, q(x)) : x \in X\}$$

be the generalized Nehari solution set. Then Theorem 2.2 states that

$$\min_{u \in \mathcal{M}} J(u) = \min_{x \in X} J(x, q(x)) \tag{3.4}$$

yields a solution $u^* = (x^*, q(x^*))$ with LRI $(u^*) = k$, which can be numerically approximated by, e.g., a steepest descent method. Since the minimization problem (3.4) involves a composite function $J(\cdot, q(\cdot))$, to do so, usually q need be C^1. However, due to the special structure of q in our approach, by the analysis in Theorem 2.1, in our steepest descent algorithm, q need only be continuous.

4 A local min-equilibrium algorithm and its convergence

In this section, we present a steepest descent algorithm to find a local minimum point of $J(\cdot, q(\cdot))$ and prove its convergence. Let $k = 0, 1, ...$ be the LRI of a solution to be found. A minimax solution of J with LRI=0 can be computed by many algorithm available in the literature, e.g., the algorithm presented in [15]. Thus we present the algorithm for finding a min-equilibrium solution with LRI$= k = 1, 2, ...$

4.1 The flow chart of the algorithm

Step 1: Given $\varepsilon > 0, \lambda > 0$. Let $x_0 \in X$ be an initial guess for a solution x^* and initial guesses $\bar{y}_0^0, \bar{y}_0^1, ..., \bar{y}_0^k$. Let $n = 0$;

Step 2: Using the minimax algorithm developed in [4, 5] to find saddle points $y_n^0, y_n^1, ..., y_n^{k-1}$ of $J(x_n, \cdot)$ recursively, i.e., a saddle point y_n^i of $J(x_n, \cdot)$ with LSI$= i$ is found by the minimax algorithm. In more details

Substep 2.1: using the initial guess $\bar{y}_n^0$ to solve for

$$y_n^0 = \arg\max_{y \in Y} J(x_n, y);$$

Substep 2.2: for $i = 1, ..., k$, let $L_{x_n} = [0, y_n^1, ..., y_n^{i-1}]$ be spanned by the previously found critical points $y_n^1, ..., y_n^{i-1}$ of $J(x_n, y_n^0 + \cdot)$ and using the initial guess $\bar{y}_n^i$ to solve for

$$y_n^i = \arg \max_{y \in S_{L_{x_n}^\perp}} \min_{z \in [L_{x_n}, y]} J(x_n, y_n^0 + z).$$

Substep 2.3: Denote the solution by $q(x_n) = y_n^0 + y_n^k$;

Step 3: Compute the steepest descent vector $d_n = J'_x(x_n, q(x_n))$;

Step 4: If $\|d_n\| \le \varepsilon$ then output $(x_n, q(x_n))$, stop; else goto Step 5;

Step 5: Set $x_n(s) = x_n - sd_n$ and find the stepsize

$$s_n = \max_{m \in \mathbb{N}} \left\{ \frac{\lambda}{2^m} \,\middle|\, 2^m > \|d_n\|, J(x_n(\frac{\lambda}{2^m}), q(x_n(\frac{\lambda}{2^m}))) - J(x_n, q(x_n)) \le -\frac{\lambda}{2^{m+1}} \|d_n\|^2 \right\}$$

where initial guesses $\bar{y}_0^0, \bar{y}_0^1, ..., \bar{y}_0^k$ are used to find $q(x_n(\frac{\lambda}{2^m}))$ as stated in Step 2.

Step 6: Set $x_{n+1} = x_n(s_n), \bar{y}_{n+1}^i = y_n^i, 0 \le i \le k$ and update $n = n + 1$ then goto **Step 2**.

It is very important to use the initial guesses stated in the algorithm in order to consistently trace and keep in the same y-equilibrium selection q and to make q continuous.

4.2 Some convergence results

Due to the composition of the function and multiplicity of solutions, convergence analysis of the algorithm becomes extremely complicated. We assume that q is continuous.

For $x \in X, d = J'_x(x, q(x)) \ne 0$, define $x(s) = x - sd$ and define the exact stepsize

$$s(x) = \max_{\lambda \ge s > 0} \left\{ s \,\middle|\, \lambda \ge s\|d\|, J(x(s), q(x(s))) - J(x, q(x)) \le -\frac{s}{2}\|d\|^2 \right\}.$$

Then it is clear from Step 5 in the algorithm that

$$\frac{1}{2}s(x_n) \le s_n \le s(x_n).$$

THEOREM 4.1. *If $x_0 \in X, J'_x(x_0, q(x_0)) \ne 0$, then there are $s_0 > 0$ and a neighborhood $\mathcal{N}(x_0)$ of x_0 s.t. $s(x) > s_0, \forall x \in \mathcal{N}(x_0)$.*

Proof. There is $s_1 > 0$ s.t. when $s_1 > s > 0$, we have $\lambda > s\|J'_x(x_0(s), q(x_0(s)))\|$ and

$$J(x_0(s), q(x_0(s))) - J(x_0, q(x_0)) < -\frac{1}{2}s\|J'_x(x_0, q(x_0))\|^2$$

or

$$\frac{J(x_0(s), q(x_0(s))) - J(x_0, q(x_0))}{s\|J'_x(x_0, q(x_0))\|^2} < -\frac{1}{2}.$$

Fix $s_0 : s_1 > s_0 > 0$ and let x_0 vary, the term

$$\frac{J(x(s_0), q(x(s_0))) - J(x, q(x))}{s_0\|J'_x(x, q(x))\|^2}$$

is continuous in x. Thus there is $\mathcal{N}(x_0)$ s.t. for each $x \in \mathcal{N}(x_0)$, we have

$$\frac{J(x(s_0), q(x(s_0))) - J(x, q(x))}{s_0\|J'_x(x, q(x))\|^2} < -\frac{1}{2},$$

i.e., $s(x) > s_0, \forall x \in \mathcal{N}(x_0)$. ∎

THEOREM 4.2. *If* $\lim_{n\to\infty} J(x_n, q(x_n)) > -\infty$ *where* $\{(x_n, q(x_n))\}$ *is the sequence generated by the algorithm, then there exists a subsequence* $\{x_{n_i}\} \subset \{x_n\}$ *s.t.* $\nabla J(x_{n_i}, p(x_{n_i})) \to 0$ *as* $i \to \infty$.

Proof. Suppose not. Then there is $\delta > 0$ s.t. $\|J'_x(x_n, q(x_n))\|_X = \|\nabla J(x_n, q(x_n))\|_H > \delta$ when n is large, say $n > n_0$. By our stepsize rule, we have

$$\begin{aligned} J(x_{n+1}, q(x_{n+1})) - J(x_n, q(x_n)) &\leq -\frac{1}{2}s(x_n)\|J'_x(x_n, q(x_n))\|^2 \\ &= -\frac{1}{2}\|x_{n+1} - x_n\|\|J'_x(x_n, q(x_n))\|. \end{aligned}$$

Adding up two sides from n_0 to ∞, we have

$$\begin{aligned} -\infty &< \lim_{n\to\infty} J(x_n, q(x_n)) - J(x_{n_0}, q(x_{n_0})) = \sum_{n=n_0}^{\infty} J(x_{n+1}, q(x_{n+1})) - J(x_n, q(x_n)) \\ &\leq -\frac{1}{2}\sum_{n=n_0}^{\infty} \|x_{n+1} - x_n\|\|J'_x(x_n, q(x_n))\| \leq -\frac{1}{2}\delta\sum_{n=n_0}^{\infty} \|x_{n+1} - x_n\|, \end{aligned}$$

which implies that $\{x_n\}$ is a Cauchy sequence. Thus there is $x^* \in X$ s.t. $x_n \to x^*$ as $n \to \infty$. Since $J'_x(x, q(x))$ is continuous in x, we have $\|\nabla J(x^*, q(x^*))\|_H = \|J'_x(x^*, q(x^*))\|_X \geq \delta$. By Theorem 4.1, there are $s_0 > 0$ and a neighborhood $\mathcal{N}(x^*)$ s.t. $s(x) \geq s_0, \forall x \in \mathcal{N}(x^*)$. This implies that when n is large, $x_n \in \mathcal{N}(x^*)$ and $\|x_{n+1} - x_n\| = s_n\|J'_x(x_n, q(x_n))\| \geq \frac{1}{2}s_0\delta$. Then $\{x_n\}$ is not a Cauchy sequence. It leads to a contradiction. ∎

Note that by the definition of q, $J(x_n, q(x_n))$ has a much better chance to be bounded from below. To establish more convergence results, we need the following compactness assumption which is commonly used in the literature.

Definition 4.1. *A C^1-function $J : X \times Y \to \mathbb{R}$ is said to satisfy PS (Palais-Smale) condition if any sequence $\{(x_n, y_n)\} \subset X \times Y$ s.t. $J(x_n, y_n)$ is bounded and $\nabla J(x_n, y_n) \to 0$ has a convergent subsequence.*

From now on we assume J satisfies PS condition. Note that $J'_y(x, q(x)) = 0$. If $J(x, q(x))$ is bounded from below where q is a continuous y-equilibrium selection of J in Y, then Ekeland's variational principle guarantees the existence of a solution x^* to

$$\min_{x\in X} G(x) = \min_{x\in X} \max_{y\in Y} J(x, y).$$

Let $K = \{x \in X : J'_x(x, q(x)) = 0\}$ and $K_c = \{x \in K : J(x, q(x)) = c\}$. Then by PS condition, the set K_c is always compact. By adopting a similar approach in [11], we prove the following point-to-set convergence result.

Theorem 4.3. *Given two open sets $\emptyset \neq V_2 \subset V_1 \subset X$. Assume (a) $V_1 \cap K \subset K_c$ where $c = \inf_{x\in V_1} J(x, q(x)) > -\infty$ and (b) there is $d > 0$ s.t.*

$$\inf\{J(x, q(x)) : x \in V_1, dis(x, \partial V_1) \leq d\} = a > b = \sup\{J(x, q(x)) : x \in V_2\} > c.$$

Let $\lambda < d$ be given in the algorithm and $\{x_n\}$ be the sequence generated by the algorithm started from an initial guess $x_0 \in V_2$. Then $\forall \varepsilon > 0$, there is $N > 0$ s.t.

$$dis(x_n, K_c) < \varepsilon, \quad \forall n > N.$$

Proof. Let $G(x) \equiv J(x, q(x))$. Since $\lambda < d$, by our stepsize rule, if $x_n \in V_1$ and $\text{dis}(x_n, \partial V_1) > d$, then $x_{n+1} \in V_1$. By the monotone decreasing nature of our algorithm and (b), $x_0 \in V_2$ implies $G(x_1) < b, x_1 \in V_1$. Thus $x_2 \in V_1$. Then $G(x_n) < b$ and $x_n \in V_1$ imply $\text{dis}(x_n, \partial V_1) > d$ and $x_{n+1} \in V_1$, i.e., $\{x_n\} \subset V_1$. Theorem 4.2 states that $\{x_n\}$ has a subsequence that converges to $x^* = \arg\min_{x\in X} G(x)$. By (a) and Theorem 2.2, we must have $x^* \in \bar{V_1} \cap K_c \neq \emptyset$. By the monotone decreasing nature of our algorithm, we have $\lim_{n\to\infty} G(x_n) = c$. Denote

$$\bar{G}(x) = \begin{cases} G(x), & x \in \bar{V_1}, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is clear that $\bar{G}$ is lower semi-continuous on X. By Ekeland's variational principle, for each $\delta_n = (G(x_n) - c)^{\frac{1}{2}}$, there is $\bar{x}_n \in X$ s.t.

$$\begin{aligned} \bar{G}(\bar{x}_n) - \bar{G}(x) &\leq \delta_n \|\bar{x}_n - x\|, \quad \forall x \in X, \\ \bar{G}(\bar{x}_n) - G(x_n) &\leq -\delta_n \|\bar{x}_n - x_n\|. \end{aligned}$$

It is clear that $\bar{x}_n \in \bar{V}_1$. Then we have

$$\begin{aligned} G(\bar{x}_n) - G(x) &\leq \delta_n \|\bar{x}_n - x\|, \quad \forall x \in \bar{V}_1, \\ G(\bar{x}_n) - G(x_n) &\leq -\delta_n \|\bar{x}_n - x_n\|. \end{aligned} \tag{4.1}$$

It follows $c \leq G(\bar{x}_n) \leq G(x_n) - \delta_n \|\bar{x}_n - x_n\|$ or $\|\bar{x}_n - x_n\| \leq \delta_n^{\frac{1}{2}}$. Then $G(x_n) \to c$ implies $\delta_n \to 0$ and $G(\bar{x}_n) \to c$. For those large n with $J'_x(x_n, q(x_n)) \neq 0$, by Theorem 2.1, when $s > 0$ is small, we have

$$G(\bar{x}_n(s)) - G(\bar{x}_n) \leq -\frac{1}{2} \|\bar{x}_n(s) - \bar{x}_n\| \|J'_x(\bar{x}, q(\bar{x}_n))\|.$$

Hence by (4.1)

$$\|\nabla J(\bar{x}_n, q(\bar{x}_n))\|_H = \|J'_x(\bar{x}_n, q(\bar{x}_n))\|_X \leq 2\delta_n \to 0,$$

which also implies

$$\|\nabla J(x_n, q(x_n))\| \to 0. \tag{4.2}$$

Let β be any limit point of $\{\text{dis}(x_n, K_c)\}$ and $\{x_{n_k}\} \subset \{x_n\}$ be a subsequence s.t. $\beta = \lim_{k\to\infty} \text{dis}(x_{n_k}, K_c)$. Since $J(x_{n_k}, q(x_{n_k})) \to c$ and $\nabla J(x_{n_k}, q(x_{n_k})) \to 0$ as $k \to \infty$, by PS condition, $\{x_{n_k}\}$ has a subsequence, denoted by $\{x_{n_k}\}$ again, s.t. $x_{n_k} \to x^* \in \bar{V}_1$. Then (4.2) leads to $x^* \in K_c$ or $\beta = 0$. ∎

Condition (b) in Theorem 4.3 is similar to those in linking theorems in critical point theory to form a trap [7,9]. Theorem 4.3 directly leads to a usual point-to-point convergent result if $V_1 \cap K_c$ contains only one point.

5 Differentiability of a Trough Selection p

Let $u^* = (x^*, y^*)$ be a min-equilibrium state of J as defined in Theorem 3.2 and p be a trough selection. To determine LRI of u^* by Theorem 3.2, we need the condition that a trough selection p is differentiable at y^*. How to check it.

Assume $L = [w_1, w_2, .., w_n] \subset Y$. First we generalize the trough mapping P and a trough selection p to the L-$\perp$ mapping P and an L-$\perp$ selection p by defining, for each $y \in S_{L^\perp}$,

$$P(y) = \{ty + y_L : y_L \in L, J'_y(x^*, ty + y_L) \perp [L, y]\} \quad \text{and} \quad p(y) \in P(y).$$

It is clear that the trough mapping satisfies the orthogonal condition in the definition of the L-$\perp$ mapping and a trough selection is an L-$\perp$ selection. This generalization enables us to discuss a limit of y and also to apply the implicit function theorem.

By the definition, an L-$\perp$ selection $p(y^*) = t_0 y^* + t_1 w_1 + ... + t_n w_n$ is solved from (n+1) orthogonal conditions

$$\begin{aligned} F_0(y^*, t_0, t_1, ..., t_n) &\equiv \langle J'_y(x^*, t_0 y^* + t_1 w_1 + ... + t_n w_n), y^* \rangle = 0, \\ F_j(y^*, t_0, t_1, ..., t_n) &\equiv \langle J'_y(x^*, t_0 y^* + t_1 w_1 + ... + t_n w_n), w_j \rangle = 0, \ j = 1, ..., n. \end{aligned}$$

Thus we have a system of $(n+1) \times (n+1)$ terms

$$\begin{aligned} \frac{\partial F_0}{\partial t_0} &= \langle J''_{yy}(x^*, t_0 y^* + t_1 w_1 + ... + t_n w_n) y^*, y^* \rangle, \\ \frac{\partial F_0}{\partial t_i} &= \langle J''_{yy}(x^*, t_0 y^* + t_1 w_1 + ... + t_n w_n) w_i, y^* \rangle, \quad i = 1, 2, ..., n. \\ \frac{\partial F_j}{\partial t_0} &= \langle J''_{yy}(x^*, t_0 y^* + t_1 w_1 + ... + t_n w_n) y^*, w_j \rangle, \quad j = 1, 2, ..., n. \\ \frac{\partial F_j}{\partial t_i} &= \langle J''_{yy}(x^*, t_0 y^* + t_1 w_1 + ... + t_n w_n) w_i, w_j \rangle, \quad i, j = 1, 2, ..., n. \end{aligned}$$

By the implicit function theorem, if the $(n+1) \times (n+1)$ matrix

$$(5.1) \quad Q'' = \begin{bmatrix} \langle J''_{yy}(x^*, p(y^*)) y^*, y^* \rangle, & \langle J''_{yy}(x^*, p(y^*)) w_1, y^* \rangle, & ... & \langle J''_{yy}(x^*, p(y^*)) w_n, y^* \rangle \\ \langle J''(x^*, p(y^*)) y^*, w_1 \rangle, & \langle J''_{yy}(x^*, p(y^*)) w_1, w_1 \rangle, & ... & \langle J''_{yy}(x^*, p(y^*)) w_n, w_1 \rangle \\ ... & & ... & \\ \langle J''_{yy}(x^*, p(y^*)) y^*, w_n \rangle, & \langle J''_{yy}(x^*, p(y^*)) w_1, w_n \rangle, & ... & \langle J''_{yy}(x^*, p(y^*)) w_n, w_n \rangle \end{bmatrix}$$

is invertible or $|Q''| \neq 0$, where $p(y^*) = t_0 y^* + t_1 w_1 + ... + t_n w_n$, then p is differentiable at and near y^*. This condition can be easily and numerically checked.

References

[1] J.J. Garcia-Ripoll, V.M. Perez-Garcia, E.A. Ostrovskaya and Y. S. Kivshar, "Dipole-mode vector solitons", *Phy. Rev. Lett.* 85(2000) 82-85.

[2] J.J. Garcia-Ripoll and V.M. Perez-Garcia, "Optimizing Schrodinger functionals using Sobolev gradients: Applications to quantum mechanics and nonlinear optics", *SIAM Sci. Comp.* 23(2001) 1316-1334.

[3] G. Henkelman, G. Johannesson and H. Jonsson, Methods for finding saddle points and minimum energy paths, *Comput. Chemistry*, D. Schwartz, ed., vol. 5, Kluwer, 2000.

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

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

[6] Z.H. Musslimani, M. Segev, D.N. Christodoulides and M. Soljacic, Composite multihump vector solitons carrying topological charge, *Phy. Rev. Lett.* 84(2000), 1164-1167.

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

[8] J. Smoller, *Shack Waves and Reaction-Diffusion Equatios*, Springer-Verlag, New York, 1983.

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

[10] X. Yao and J. Zhou, A minimax method for finding multiple critical points in Banach spaces and its application to quasilinear elliptic PDE, *SIAM J. Sci. Comp.*, 26(2005), 1796-1809.

[11] X. Yao and J. Zhou, A unified convergence result on a minimax method for finding multiple critical points in Banach spaces, preprint.

[12] J. Zhou, A local min-orthogonal method for finding multiple saddle points, *J. Math. Anal. Appl.*, 291 (2004), 66-81.

[13] J. Zhou, Instability analysis of saddle points by a local minimax method, *Math. Comp.*, 74(2005), 1391-1411.

[14] J. Zhou and X. Yao, Stabilization in computing saddle points. *Control and Boundary Analysis*, Marcel Dekker, New York, 2004.

[15] J. Zhou, Saddle critical point analysis and computation, *Nonlinear Analysis*, 63(2005), 1000-1009.