Vietnam Journal of Mathematics **40**:2&3(2012) 231-253

Vietnam Journal of MATHEMATICS

Gap Functions and Error Bounds for Variational and Generalized Variational Inequalities

Joydeep Dutta

Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur-208016, India

Dedicated to my friend Professor Phan Quoc Khanh on his 65th birthday

Received May 02, 2012
Revised September 11, 2012

Abstract. It is well known that gap functions play a major role in developing error bounds for strongly monotone and affine monotone variational inequalities. However the error bound would be important from the point of view of computation if the gap function values diminish on sequences of points which converge to a solution of the variational inequality. A gap function exhibiting such behavior is termed as a well-behaved gap function. In this article we would like to study some important gap functions for both variational inequalities with single-valued maps and set-valued maps and see what additional conditions one must impose on them so that they are well behaved.

2000 Mathematics Subject Classification. 90C33.

Key words. Variational inequalities, generalized variational inequalities, gap functions, error bounds, strong monotonicity, strong convexity.

1. Introduction and motivation

Variational inequalities in finite dimensions generalize the necessary and sufficient optimality condition for minimizing a differentiable convex function over a convex set. More formally if $F : \mathbb{R}^n \to \mathbb{R}^n$ and let C be a convex set then the variational inequality problem $VI(F, C)$ consists of finding $\bar{x} \in C$ such that

$$\langle F(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \forall x \in C.$$

It is simple to observe that if $F(x) = \nabla f(x)$ for all $x \in \mathbb{R}^n$, where f is a differentiable convex function then the expression in the variational inequality is nothing but the necessary and sufficient condition for minimizing the convex function f over the convex set C. More generally if we consider f to be non-differentiable then the generalization of the necessary and sufficient optimality condition leads to the so-called *generalized variational inequality* or GVI for short. Given a set-valued map $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and a convex set C the generalized variational inequality $GVI(T, C)$ consists of finding $\bar{x} \in C$ and $\xi \in T(\bar{x})$ such that

$$\langle \xi, x - \bar{x} \rangle \geq 0 \quad \forall x \in C.$$

It is simple to observe that if $T = \partial f$ where ∂f represents the subdifferential multifunction of the convex function f then the above GVI is nothing but the non-smooth necessary and sufficient optimality condition for minimizing the convex function over the convex set C. However in the set-valued setting we have a flexibility and we can define what we call a weak GVI denoted as $WGVI(T, C)$, where one needs to find $\bar{x} \in C$ such that for each $x \in C$ there exists $\xi_x \in T(\bar{x})$ (depending on x) such that

$$\langle \xi_x, x - \bar{x} \rangle \geq 0.$$

The above type of variational inequalities has been referred in the literature as the *Stampacchia type variational inequalities*. For more details on the problem $VI(F, C)$, see for example Facchinei and Pang [9] and for recent results on $GVI(T, C)$ and $WGVI(T, C)$ see Aussel and Dutta [4]. There are various ways to approach a variational inequality. One approach is to devise an optimization problem whose solution would lead to the solution of the original variational inequality. This is done by introducing the notion of a *merit function* or *gap function*. Given any of the above variational inequalities a gap function is a function $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ which satisfies the following two properties

i) $\psi(x) \geq 0$ for all $x \in C$;
ii) $\psi(\bar{x}) = 0$, $\bar{x} \in C$ if and only if $\bar{x}$ solves the variational inequality.

Thus it is clear that if we minimize the function ψ over the convex set C and if $\bar{x}$ is the minimizer with $\psi(\bar{x}) = 0$ then $\bar{x}$ is a solution of the variational inequality. However as one might observe that it is not so simple to minimize the function ψ since it could not only be non-convex but every value of the function ψ is evaluated by solving an optimization problem. Thus it becomes quite cumbersome to minimize the function ψ. On the other hand it has been observed that gap functions play a fundamental role for developing error bounds for certain classes of monotone variational inequalities. An *error bound* is an expression which provides an upper estimate of the distance of an arbitrary feasible point to the solution set of the variational inequality. For more details see for example Pang [14] and Facchinei and Pang [9]. Let S denote the solution set of a variational inequality. An error bound is given by the following expression

$$d(x, S) \leq k\beta(\psi(x)),$$

where $x \in C$ and β is a real-valued function on $\mathbb{R}$ such that $\beta(0) = 0$ and $k > 0$ is a constant. Such types of error bounds are called global error bounds. In this article we will consider only global error bounds. In fact all the error bounds appearing in this article will be of the form

$$d(x, S) \leq k\sqrt{\psi(x)}, \quad x \in C.$$

In order for the above error bounds to be meaningful it is important to note that first the gap function must be finite valued and when x is indeed very near the solution set then the gap function value at x must also be very near zero. We call a gap function **well-behaved** if for any sequence $\{x_k\}$ with $x_k \in C$ such that $x_k \to \bar{x}$ where $\bar{x}$ is a solution of $VI(F, C)$ or $GVI(T, C)$ we have $\psi(x_k) \to 0$ as $k \to \infty$.

All the gap functions for $VI(F, C)$ that we discuss in this article are well-behaved since most of them are continuous under very mild and natural assumptions. In the last section of the article we consider $GVI(T, C)$ and we will see that the gap function is well-behaved under slightly stringent assumptions.

The article is arranged as follows. The second section discusses the well known Auslender gap function and a gap function for hemivariational inequalities. For hemivariational inequalities an error bound is developed under more natural conditions than that given by Chen, Goh and Yang [8]. Further when F is monotone and continuous and C is a closed convex set one can construct a gap function for $VI(F, C)$ which is always convex. This is known as the dual gap function. The dual gap function is always a convex function. However it is quite difficult to compute the values of the dual gap function.

In Section 3 we focus on the primal generalized gap function introduced by Auchmuty [1]. The primal generalized gap function is a versatile one but did not receive much attention in the literature. The flexibility of the generalized gap function arises from the fact that a proper lower-semicontinuous convex function is used to define it. By choosing particular convex functions we can recover some important class of gap functions like the regularized gap function in the sense of Fukushima [10]. We derive some error bounds using a class of regularized gap functions in the sense of Fukushima [10]. It appears that the error bound we present here does not seem to have been explicitly mentioned in the literature and to the best of our efforts we could not find it. We also discuss how the error bound using the primal generalized gap function as given by Larsson and Patriksson [11] can be modified. Thus Section 2 and Section 3 have some pedagogical flavor.

In Section 4 we introduce a new gap function for the generalized variational inequality $GVI(T, C)$. Our gap function for $GVI(T, C)$ is motivated by the gap function for $WGVI(T, C)$ introduced by Aussel and Dutta [4]. Further we develop a regularized gap function for $GVI(T, C)$ and use it to devise an error bound for $GVI(T, C)$ when T is a strongly monotone map. We end the article by discussing the conditions which make the regularized gap function for $GVI(T, C)$ well-behaved.

We shall end this section by mentioning the definitions of monotonicity and strong monotonicity which will be used in the sequel.

A function $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be monotone if for any $x, y \in \mathbb{R}^n$,

$$\langle F(y) - F(x), y - x \rangle \geq 0.$$

The function F is called strongly monotone with strong monotonicity parameter $\mu > 0$ if for any $x, y \in \mathbb{R}^n$

$$\langle F(y) - F(x), y - x \rangle \geq \mu \|y - x\|^2.$$

An important example of a strongly monotone operator is $F(x) = Mx + q$ where M is a $n \times n$ symmetric positive definite matrix with $\mu = \lambda_{\min}(M)$ where $\lambda_{\min}(M)$ denotes the minimum eigenvalue of A. We shall usually refer a strongly monotone map with strong monotonicity parameter $\mu > 0$ as a μ-strongly monotone map.

Let $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued map. The map T is said to be monotone if for any $x, y \in \mathbb{R}^n$ and for any $x^* \in T(x)$ and $y^* \in T(y)$ we have

$$\langle y^* - x^*, y - x \rangle \geq 0.$$

And T is said to be μ-strongly monotone if there exists $\mu > 0$ such that for any $x, y \in \mathbb{R}^n$ and any $x^* \in T(x)$ and $y^* \in T(y)$ we have

$$\langle y^* - x^*, y - x \rangle \geq \mu \|y - x\|^2.$$

In the rest of the paper we shall consider C to be a closed convex set.

2. Auslender gap function

In the literature on variational inequalities it is usually mentioned that the gap function referred to Auslender appeared in his book on numerical optimization [2]. However to the best of our knowledge it seems that it first appeared in [3].

The Auslender gap function denoted by θ is defined as

$$\theta(x) = \sup_{y \in C} \langle F(x), y - x \rangle.$$

First of all let us note that in order to calculate the value of θ at a given value x we need to minimize an affine function in y over the convex set C. It is important to see that θ is an extended valued function. Of course we need additional properties on C so that θ becomes finite. One such property is that C is compact. If F is a continuous function then it is quite simple to see that θ is a lower-semicontinuous function. Thus at least we need to have F to be continuous. The following result due to Marcotte [12], tells under what condition the function θ can be Lipschitz over the set C.

Theorem 2.1. *Let C be a compact convex set and F be a continuously differentiable map. Then θ is Lipschitz continuous on C, i.e., there exists a constant $l \geq 0$ such that*

$$|\theta(x) - \theta(z)| \leq l\|x - z\|.$$

An important example for such an F is the affine function $F(x) = Mx + q$ where M is a $n \times n$ symmetric matrix and $q \in \mathbb{R}^n$. When M is symmetric and positive semidefinite then it is easy to see that θ is a convex function. If C is compact then the subdifferential of θ can be computed very easily using the Danskin's formula. When C is compact and F is μ-strongly monotone then it is simple to show that for any $x \in C$ then we have

$$\|x - \bar{x}\| \leq \sqrt{\frac{\theta(x)}{\mu}},$$

where $\bar{x}$ is the unique solution of $VI(F, C)$.

Unfortunately even if C is polyhedral there is no way to guarantee that θ will be finite. The most well known case is the one where $C = \mathbb{R}^n_+$. Then we have the famous *nonlinear complementarity problem* in which we have to find $x \in \mathbb{R}^n_+$ such that $F(x) \in \mathbb{R}^n_+$ and $\langle F(x), x \rangle = 0$. In this case the function $\theta(x)$ is given as follows. If $F(x) \in \mathbb{R}^n_+$ then $\theta(x) = \langle x, F(x) \rangle$ and if $F(x) \notin \mathbb{R}^n_+$ then $\theta(x) = +\infty$.

There is another class of variational inequalities called *Hemivariational Inequalities* (HVI for short) is also studied in the literature. Given $F : \mathbb{R}^n \to \mathbb{R}^n$ and a convex function $f : \mathbb{R}^n \to \mathbb{R}$ the problem $HVI(F, f, C)$ consists in finding $\bar{x} \in C$ such that

$$\langle F(\bar{x}), y - \bar{x} \rangle + f(y) - f(\bar{x}) \geq 0 \quad \forall y \in C.$$

Note that if f is a differentiable convex function then $HVI(F, f, C)$ is equivalent to $VI(F + \nabla f, C)$. If f is not differentiable then $H(F, f, C)$ would correspond to a $GVI(F + \partial f, C)$. This means that if $\bar{x}$ solves $HVI(F, f, C)$ then there exists $\xi \in \partial f(\bar{x})$ such that

$$\langle F(\bar{x}) + \xi, y - \bar{x} \rangle \geq 0 \quad \forall y \in C.$$

Conversely if $\bar{x}$ solves $GVI(F(\bar{x}) + \partial f, C)$ then it solves $HVI(F, f, C)$. These simple facts will be crucial in devising the error bounds for $HVI(F, f, C)$. In Chen, Goh and Yang [8] the function f is considered to be an extended valued proper lower-semicontinuous convex function. They call such an HVI an *extended variational inequality*. As proposed by Chen, Goh and Yang [8] the gap function for $HVI(F, f, C)$ is defined as follows

$$\phi(x) = \sup_{y \in C} \{\langle F(x), x - y\rangle + f(x) - f(y)\}.$$

Note that just like the Auslender gap function in general it is not possible to say whether the above function ϕ is finite or not unless C has some additional properties like compactness. However if f is a strongly convex function and C is a closed convex set then ϕ is finite without any boundedness assumption on the set C. In fact observe that one can write

$$\phi(x) = -\inf_{y \in C} \{\langle F(x), y - x\rangle + f(y) - f(x)\}.$$

Note that for each fixed $x \in \mathbb{R}^n$ the function

$$y \mapsto \langle F(x), y - x\rangle + f(y) - f(x)$$

is strongly convex and thus has a unique minimizer on the closed convex set C. It is simple enough to show that ϕ is a gap function for $HVI(F, f, C)$. Let us note that if F is strongly monotone or f is strongly convex, $HVI(F, f, C)$ has a unique solution. This can be seen easily by considering the equivalence of $HVI(F, f, C)$ with $GVI(F + \partial f, C)$.

Theorem 2.2. *Let us consider $HVI(F, f, C)$ and let $\bar{x}$ be the unique solution of $HVI(F, f, C)$, then we have the following*

i) *Let C be compact and F be μ-strongly monotone then for any $x \in C$*

$$\|x - \bar{x}\| \leq \sqrt{\frac{\phi(x)}{\mu}}.$$

ii) *Let F be monotone and f be ρ-strongly convex $(\rho > 0)$ then*

$$\|x - \bar{x}\| \leq \sqrt{\frac{\phi(x)}{\rho}}.$$

Proof. Let us consider the case i) where we assume that F is μ-strongly monotone. Now from the definition of the gap function ϕ we have for all $y \in C$

$$\phi(x) \geq \langle F(x), x - y\rangle + f(x) - f(y).$$

In particular for $y = \bar{x}$ we have

$$\phi(x) \geq \langle F(x), x - \bar{x}\rangle + f(x) - f(\bar{x}).$$

Now using the strong monotonicity of F and the convexity of f for any $\xi \in \partial f(\bar{x})$ we have

$$\phi(x) \geq \langle F(\bar{x}), x - \bar{x} \rangle + \mu \|x - \bar{x}\|^2 + \langle \xi, x - \bar{x} \rangle.$$

Since $\bar{x}$ is the unique solution of $HVI(F, f, C)$ it is also a solution of $GVI(F + \partial f, C)$ there exists $\bar{\xi} \in \partial f(\bar{x})$ such that

$$\langle F(\bar{x}) + \bar{\xi}, x - \bar{x} \rangle \geq 0.$$

Hence we have that

$$\phi(x) \geq \mu \|x - \bar{x}\|^2.$$

This establishes the error bound in the first case. In the second case we have assumed that f is strongly convex. Hence we know from Vial [17] that for all $\xi \in \partial f(\bar{x})$

$$f(x) - f(\bar{x}) \geq \langle \xi, x - \bar{x} \rangle + \rho \|x - \bar{x}\|^2.$$

Using this fact and the monotonicity of F we conclude that for all $\xi \in \partial f(\bar{x})$ we have

$$\phi(x) \geq \langle F(\bar{x}), x - \bar{x} \rangle + \langle \xi, x - \bar{x} \rangle + \rho \|x - \bar{x}\|^2.$$

The rest of the proof follows as the first case by noting that $\bar{x}$ also solves $GVI(F + \partial f, C)$. This completes the proof. ■

It is important to note that we have considered f to be a finite valued convex function on $\mathbb{R}^n$ while Chen, Goh and Yang [8] and Patriksson [15] consider f to be an extended-valued proper lower-semicontinous convex function. However it is important to note that when f is proper and lower-semicontinuous and if $\bar{x}$ is the solution then necessarily $f(\bar{x})$ has to be finite. Thus without loss of generality we can always consider f to be finite. This also makes sense from the computational point of view. Further when f is an extended-valued, proper and lower-semicontinuous convex function and $\text{int}(\text{dom}\, f) \cap C \neq \emptyset$ it has been shown in Patriksson [15] $HVI(F, f, C)$ is equivalent to $GVI(F + \partial f, C)$. Further it is important to note that even if $\bar{x} \in \text{dom}\, f$ there is no guarantee that $\partial f(\bar{x}) \neq \emptyset$ unless $\bar{x} \in \text{int}(\text{dom}\, f)$. Note that if $\partial f(\bar{x})$ is empty then the inequality associated with $GVI(F + \partial f, C)$ has no meaning. Further in order to develop an error bound for $GVI(F + \partial f, C)$ the approach taken by Patriksson [15] closely follows the line of treatment of gap functions by Auchmuty [1]. On the other hand in Chen, Goh and Yang [8] under the assumption that F is μ-strongly pseudomonotone with respect to f and some further additional assumptions an error bound is developed for $HVI(F, f, C)$. See [8] for details. Let us also note that our proof of Theorem 2.2 is based on very simple and natural assumptions.

A natural question is whether we can have a gap function for $VI(F, C)$ which is always convex under natural assumptions, for example F being continuous and monotone. The answer surprisingly turns out to be "yes". We shall denote such a gap function by θ_D and define it as

$$\theta_D(x) = \sup_{y \in C} \langle F(y), x - y \rangle.$$

It is interesting to note that θ_D is always convex and is a gap function for $VI(F, C)$ when F is monotone and continuous. Note that even when $F(x) = Mx + q$, with M symmetric and positive definite it is not possible to guarantee that the Auslender gap is finite unless C is compact. Consider the case $C = \mathbb{R}^n_+$ to see this. While if $F(x) = Mx + q$ with M symmetric and positive definite then we will show that θ_D is always finite irrespective of whether C is compact or not. Note that when $F(x) = Mx + q$ we have

$$\theta_D(x) = \sup_{y \in C} \{-\langle y, My \rangle - \langle q, y \rangle + \langle Mx, y \rangle + \langle q, x \rangle\}.$$

Note that the function $y \mapsto \langle y, My \rangle + \langle q, y \rangle - \langle Mx, y \rangle - \langle q, x \rangle$ is strongly convex and we assume that C is closed it has a unique minimizer over C. Let us call this as $y(x)$. Then we have

$$\theta_D(x) = \langle My(x) + q, x - y(x) \rangle.$$

Thus θ_D is finite in this particular case. Further it is simple to see that if F is strongly monotone then θ_D is finite. However it is important to note that when C is compact then θ_D can be used to compute the solution of $VI(F, C)$ with F continuous and monotone. Note that in this case we have

$$\partial \theta_D(x) = \text{co}\{F(y(x)) : y(x) \in J(x)\},$$

where $J(x) = \text{argmax}_{y \in C} \langle F(y), x - y \rangle$. Note that as always we assume that F is continuous. If $F(x) = Mx + q$ and M is symmetric and positive definite and C is compact then θ_D is differentiable since $J(x)$ is singleton and we have $\nabla \theta_D(x) = F(y(x))$, where $J(x) = \{y(x)\}$. Thus it seems that for monotone or strongly monotone problems the dual gap function is an useful device to compute the solution of $VI(F, C)$. The natural question is that whether this gap function is also amenable in devising error bounds when F is continuous and strongly monotone. It appears that it is quite a difficult matter to devise error bounds with the dual gap function. A detailed study of the dual gap function is now carried out in the general case of a GVI in Aussel and Dutta [5] where by modifying the dual gap function suitably they developed error bounds in combination with other gap functions but with sharper constraints. Aussel and Dutta [5] also show that the dual gap function can provide an error bound in the case when $F(x) = Mx + q$, with M symmetric and positive definite without any additional assumption on C.

3. The primal generalized gap function

At this point we take a detour and have a look at Auchmuty's approach to gap functions. Though Auchmuty had an elegant approach using convex analysis his approach is not so well documented in the literature. The gap function introduced by Auchmuty [1] is denoted by θ_A and is given as

$$\theta_A(x) = \langle F(x), x\rangle + \sigma_C(-F(x)),$$

where σ_C denotes the support function of the convex set C. Thus note that

$$\theta_A(x) = \langle F(x), x\rangle + \sup_{y\in C}\langle -F(x), y\rangle.$$

This shows that $\theta_A(x) = \theta(x)$. Of course Auchmuty did not stop at this and went onto develop a more general approach to the notion of a gap function. Auchmuty [1] also defines the notion of a primal generalized gap function which is given as

$$G(x) = \sup_{y\in C} \hat{L}(x, y),$$

where $\hat{L}(x, y)$ is given as follows

$$\hat{L}(x, y) = f(x) - f(y) - \langle x - y, \nabla f(x) - F(x)\rangle,$$

where f is a proper lower-semicontinuous convex function which is smooth on an open set containing C. We can thus also assume without loss of generality that f is finite and smooth on $\mathbb{R}^n$. It is not much difficult to prove that

$$G(x) = f(x) + g^*(\nabla f(x) - F(x)) + \langle x, F(x) - \nabla f(x)\rangle.$$

Note that in the above expression g^* denotes the Fenchel conjugate of g where $g = f + \delta_C$ and δ_C is the indicator function of the convex set C.

Let us now list down some important properties of the primal generalized gap function G which are stated in [11] and [1].

Theorem 3.1. *The function G is lower-semicontinuous and G is a gap function for $VI(F, C)$. Further if f is finite-valued, differentiable and strongly convex on $\mathbb{R}^n$ then G is finite and continuous.*

It is clear that if f is a constant function then $G(x) = \theta(x)$ for all $x \in \mathbb{R}^n$. Note that in the above theorem there is no boundedness assumption on C. Fukushima [10] introduced the notion of a regularized gap function which is finite for any closed convex set C. Fukushima's regularized gap function is very popular in the literature on gap functions for variational inequalities. However as we will now see that Fukushima's regularized gap function [10] is a special case of the primal generalized gap function of Auchmuty [1]. Let us consider the case when $f(x) = \frac{1}{2}\|x\|^2$, $x \in \mathbb{R}^n$. Then we have

$$L(x, y) = \frac{1}{2}\|x\|^2 - \frac{1}{2}\|y\|^2 + \langle F(x) - x, x - y\rangle.$$

Some simple calculations will show that

$$L(x,y) = \langle F(x), x-y \rangle - \frac{1}{2}\|y-x\|^2.$$

This shows that $G(x) = g(x)$ which is a simpler version of the regularized gap function of Fukushima [10]. In fact we write

$$g(x) = \sup_{y \in C} \left\{ \langle F(x), x-y \rangle - \frac{1}{2}\|y-x\|^2 \right\}.$$

Of course since our chosen $f(x) = \frac{1}{2}\|x\|^2$ is strongly convex we conclude from Theorem 3.1 that g is finite-valued and continuous. In fact it can be shown that

$$g(x) = \langle F(x), x - \operatorname{proj}_C(x - F(x)) \rangle - \frac{1}{2}\| \operatorname{proj}_C(x - F(x)) - x\|^2.$$

Let us note that in order to get a meaningful error bound when F is strongly monotone it is useful to slightly modify the definition of g by using a tuning parameter $\alpha > 0$ and define

$$g_\alpha(x) = \sup_{y \in C} \left\{ \langle F(x), x-y \rangle - \frac{\alpha}{2}\|y-x\|^2 \right\}.$$

Further it can be easily shown that when $f(x) = \frac{\alpha}{2}\|x\|^2$ then $G(x) = g_\alpha(x)$. In fact it is not much difficult to show that

$$g_\alpha(x) = \langle F(x), x - \operatorname{proj}_C(x - \frac{1}{\alpha}F(x)) \rangle - \frac{1}{2}\| \operatorname{proj}_C(x - \frac{1}{\alpha}F(x)) - x\|^2.$$

Let us give a brief explanation of how the above expression comes about. Let $y_\alpha(x)$ be the unique minimum of the strongly convex function

$$y \mapsto \langle F(x), y-x \rangle + \frac{\alpha}{2}\|y-x\|^2$$

over the closed convex set C. Then the standard optimality condition for convex programming tells us that

$$-F(x) - \alpha(y_\alpha(x) - x) \in N_C(y_\alpha(x)).$$

Noting that $N_C(y_\alpha(x))$ is a cone we have

$$\left(x - \frac{1}{\alpha}F(x)\right) - y_\alpha(x) \in N_C(y_\alpha(x)).$$

This shows that

$$y_\alpha(x) = \operatorname{proj}_C(x - \frac{1}{\alpha}F(x)).$$

Note that in order to compute $y_\alpha(x)$ one needs to compute the projection which is equivalent to solving a strongly convex optimization problem. If $C = \mathbb{R}^n_+$

then the projection is quite simple to compute and it is given by

$$y_\alpha(x) = \max\left\{x - \frac{1}{\alpha}F(x), 0\right\},$$

where the maximum is taken in the componentwise sense.

When C is described in terms of convex inequalities and affine equalities in order to compute $y_\alpha(x)$ one can use standard techniques for solving convex optimization problems. In some special cases of course, one can give an explicit formula for $y_\alpha(x)$. For example consider

$$C = \{x \in \mathbb{R}^n : Ax = b\},$$

where A is a $m \times n$ matrix of full row rank. In this case it is not difficult to show that

$$y_\alpha(x) = \left(I - A^T(AA^T)^{-1}A\right)\left(x - \frac{1}{\alpha}F(x)\right) + A^T(AA^T)^{-1}b,$$

where A^T is the transpose of the matrix A.

Let F be μ-strongly monotone and we choose $\alpha > 0$ such that $2\mu > \alpha$ and let $\bar{x}$ be the unique solution of $VI(F, C)$, then for any $x \in C$ we have

$$\|x - \bar{x}\| \leq \sqrt{\frac{2}{2\mu - \alpha}}\sqrt{g_\alpha(x)}.$$

Let us now consider the case when $f(x) = \frac{\alpha}{2}\langle x, Px\rangle$, where P is a $n \times n$ symmetric and positive definite matrix. Then we have

$$G(x) = \sup_{y \in C}\left\{\langle F(x), x - y\rangle - \frac{\alpha}{2}\langle y - x, P(y - x)\rangle\right\}.$$

When $\alpha = 1$ this coincides with the regularized gap function introduced by Fukushima [10]. In this special case we shall denote $G(x)$ as $g_\alpha^F(x)$. This function is also finite and continuous by Theorem 3.1 . However it is slightly more complicated to find out the point which minimizes the function

$$y \mapsto \langle F(x), y - x\rangle + \frac{\alpha}{2}\langle y - x, P(y - x)\rangle.$$

This is obtained through the operation of oblique projection. Observe that if $y_\alpha^F(x)$ is the unique solution of the above function then the optimality condition states that

$$-(F(x) + \alpha P(y_\alpha^F(x) - x)) \in N_C(y_\alpha^F(x)).$$

Thus we have that there exists $v \in N_C(y_\alpha^F(x))$ such that

$$-\frac{1}{\alpha}F(x) - P(y_\alpha^F(x) - x) = v.$$

Thus

$$-\frac{1}{\alpha}P^{-1}F(x) - y_\alpha^F(x) + x = P^{-1}v.$$

This shows that

$$P\left(\left(x - \frac{1}{\alpha}P^{-1}F(x)\right) - y_\alpha^F(x)\right) \in N_C(y_\alpha^F(x)).$$

The above expression is the necessary and sufficient optimality condition for the following strongly convex optimization problem whose solution is $y_\alpha^F(x)$,

$$\min \frac{1}{2}\|y - (x - \frac{1}{\alpha}P^{-1}F(x))\|_P^2, \quad \text{subject to} \quad x \in C,$$

where $\|x\|_P$ is the norm with respect to the positive definite matrix P, i.e., $\|x\|_P^2 = \langle x, Px \rangle$. The above strongly convex problem is the problem of oblique projection and we write

$$y_\alpha^F(x) = \text{proj}_{C,P}\left(x - \frac{1}{\alpha}P^{-1}F(x)\right).$$

It is amply clear that even if the oblique projection has nice properties like Lipschitz continuity it is not so simple to compute and give an explicit expression for $y_\alpha^F(x)$ even for the case when $C = \mathbb{R}_+^n$. However we will now show that for the set C given as

$$C = \{x \in \mathbb{R}^n : Ax = b\},$$

where A as before is a $m \times n$ matrix with full row rank and $b \in \mathbb{R}^m$ we can provide an explicit expression for $y_\alpha^F(x)$. Noting that for the above C we have $N_C(y) = \text{Im}A^T$ for any $y \in C$ we can conclude that there exists $\lambda \in \mathbb{R}^m$ such that

$$-\frac{1}{\alpha}P^{-1}F(x) - (y_\alpha^F(x) - x) = P^{-1}A^T\lambda.$$

This shows that

$$y_\alpha^F(x) = x - \frac{1}{\alpha}P^{-1}F(x) - P^{-1}A^T\lambda.$$

Hence we have

$$b = Ax - \frac{1}{\alpha}AP^{-1}F(x) - AP^{-1}A^T\lambda.$$

Note that since P is positive definite so is P^{-1} and since A is of full row rank it is simple to show that $AP^{-1}A^T$ is also positive definite and thus invertible. Hence we have

$$\lambda = (AP^{-1}A^T)^{-1}Ax - (AP^{-1}A^T)^{-1}b - \frac{1}{\alpha}(AP^{-1}A^T)^{-1}AP^{-1}F(x).$$

Further some simple calculations show us that

$$y_\alpha^F(x) = (I - P^{-1}A^T(AP^{-1}A^T)^{-1}A)x + \frac{1}{\alpha}(P^{-1}A^T(AP^{-1}A^T)^{-1}AP^{-1} - P^{-1})F(x) + P^{-1}A^T(AP^{-1}A^T)^{-1}b.$$

Let us now write down an error bound in terms of the gap function $g_\alpha^F(x)$ where F is μ- strongly monotone and $\mu > \frac{\alpha}{2}\|P\|$. If x^* is the unique solution of $VI(F,C)$ then

$$\|x - x^*\| \leq \sqrt{\frac{2}{2\mu - \alpha\|P\|}}\sqrt{g_\alpha^F(x)}.$$

In the above result for error bounds one observes that μ and α are related. One can ask the question whether an error bound with g_α^F can be developed without any relation between μ and α. In this case one would have to assume that F is Lipschitz over C. To the best of our efforts we were not able to locate the theorem below and its corollary in the literature though it appears that it might also be derived from the error bound proved in [18] using a generalized version of the regularized gap functions that we have mentioned here. Thus we provide the proofs.

Theorem 3.2. *Let us consider the problem $VI(F,C)$ where F is μ-strongly monotone and Lipschitz on C with Lipschitz constant $L \geq 0$. Let $x^* \in C$ be the unique solution of $VI(F,C)$. Then for any $x \in C$ we have*

$$\|x - x^*\| \leq \frac{(\alpha\|P\| + L)}{\mu}\sqrt{\frac{2}{\alpha\lambda_{\min}(P)}}\sqrt{g_\alpha^F(x)}.$$

Proof. Let $y_\alpha^F(x)$ be the unique minimizer of the strongly convex function

$$y \mapsto \langle F(x), y - x\rangle + \frac{\alpha}{2}\langle y - x, P(y - x)\rangle.$$

Then using the standard optimality conditions for convex optimization we have

$$\langle F(x) + \alpha P(y_\alpha^F(x) - x), x^* - y_\alpha^F(x)\rangle \geq 0.$$

Further as x^* solves $VI(F,C)$ we have

$$\langle F(x^*), y_\alpha^F(x) - x^*\rangle \geq 0.$$

This shows that

$$\langle F(x) - F(x^*) + \alpha P(y_\alpha^F(x) - x), y_\alpha^F(x) - x^*\rangle \leq 0.$$

A little manipulation of the above inequality will lead to the following inequality

$$\langle F(x)-F(x^*), x-x^*\rangle + \langle y_\alpha^F(x)-x, P(y_\alpha^F(x)-x)\rangle \leq -\alpha\langle P(y_\alpha^F(x)-x), x-x^*\rangle - \langle F(x)-F(x^*), y_\alpha^F(x)-x\rangle.$$

Now using the fact that P is positive definite, F is Lipschitz and μ-strongly monotone we have that

$$\mu\|x-x^*\|^2 \leq \alpha\|P\|\|y_\alpha^F(x)-x\|\|x-x^*\| + L\|x-x^*\|\|y_\alpha^F(x)-x\|.$$

This leads to the fact

$$\|x-x^*\| \leq \frac{\alpha\|P\|+L}{\mu}\|y_\alpha^F(x)-x\|. \tag{1}$$

Further observe that from the optimality conditions we have

$$\langle F(x)+\alpha P(y_\alpha^F(x)-x), x-y_\alpha^F(x)\rangle \geq 0.$$

Thus we have

$$\alpha\langle P(y_\alpha^F(x)-x), y_\alpha^F(x)-x\rangle \leq -\langle F(x), y_\alpha^F(x)-x\rangle. \tag{2}$$

Since P is positive definite we have

$$\frac{\alpha}{2}\langle y_\alpha^F(x)-x, P(y_\alpha^F(x)-x)\rangle \geq \frac{\alpha}{2}\lambda_{\min}(P)\|y_\alpha^F(x)-x\|^2.$$

This shows that

$$\alpha\langle y_\alpha^F(x)-x, P(y_\alpha^F(x)-x)\rangle - \frac{\alpha}{2}\langle y_\alpha^F(x)-x, P(y_\alpha^F(x)-x)\rangle \geq \frac{\alpha}{2}\lambda_{\min}(P)\|y_\alpha^F(x)-x\|^2.$$

Now from further manipulation using (2) along with the above inequality we obtain that

$$g_\alpha^F(x) \geq \frac{\alpha}{2}\lambda_{\min}(P)||y_\alpha^F(x)-x||^2.$$

Now the result is established by using (1). ■

Corollary 3.3. *Let us consider the problem $VI(F,C)$ with $F(x) = Mx + q$, where M is a $n \times n$ symmetric and positive definite matrix. Let x^* be the unique solution of $VI(F,C)$. Then for any $x \in C$ we have*

$$\|x-x^*\| \leq \frac{(\alpha\sqrt{\operatorname{trace} P^2} + \sqrt{\operatorname{trace} M^2})}{\lambda_{\min}(M)}\sqrt{\frac{2}{\alpha\lambda_{\min}(P)}}\sqrt{g_\alpha^F(x)}.$$

Proof. The result can be established by noting that in this case $L = \|M\|$ and $\mu = \lambda_{\min}(M)$ and then applying Theorem 3.2. Note that we use the Frobenius norm for the matrices M and P and thus $\|M\|^2 = \text{trace}(M^2)$ and $\|P\|^2 = \text{trace}(P^2)$. ■

It is important to note that in the above formula any matrix norm can be used since all matrix norms are equivalent and possibly the norm that we have used might be easy to compute in most case. For example if P is a diagonal matrix with positive entries then the above choice of the norm possibly will turn out to be quite efficient.

Let us now consider the case where $f(x) = \alpha h(x)$ with $h : \mathbb{R}^n \to \mathbb{R}$ a differentiable strongly convex function with $\mu > 0$ as the modulus of strong convexity. In this case we denote $G(x)$ as $g_\alpha^h(x)$ and we have

$$g_\alpha^h(x) = \sup_{y \in C} \{\langle F(x), x - y\rangle - \alpha D_h(y, x)\},$$

where $D_h(y, x) = h(y) - h(x) - \langle \nabla h(x), y - x\rangle$. The function $D_h(y, x)$ is called the Bregman distance induced by h which is a non-Euclidean distance and which plays a pivotal role in the generalized version of the proximal point algorithm for solving convex optimization problem and variational inequalities with maximal monotone maps. For more details see for example the recent monograph by Burachik and Iusem [7]. By arguments similar to the previous cases it is clear that g_α^h is well defined in the sense that it is finite and continuous. Thus the above gap function generalizes the regularized gap function due to Fukushima [10]. However the function g_α^h can also be defined by considering that h is just convex and not strongly convex and in such a case we can just conclude that $g_\alpha^h(x)$ is an extended-valued lower-semicontinuous function. In fact when h is just convex an error bound can be obtained in terms of g_α^h by slightly tweaking the proof of theorem of Larsson and Patriksson [11]. Though it might be difficult to verify the condition under which it holds and also the question about its finiteness remains. Thus a better approach would be use the case when h is strongly convex which will guarantee the finiteness and continuity of g_α^h and in the proof of the error bound one just need to use the fact that h is convex rather than using the more stronger property of strong convexity. Thus we will have the following result.

Theorem 3.4. *Let us consider the problem $VI(F, C)$, where F is μ-strongly monotone and the function h is strongly convex whose derivative is Lipschitz continuous of Lipschitz constant $L > 0$ over C. Let us choose $\alpha > 0$ such that $\frac{\mu}{L} > \alpha$. If $x^* \in C$ is the unique solution of $VI(F, C)$ and $x \in C$, then we have*

$$\|x - x^*\| \leq \sqrt{\frac{1}{\mu - \alpha L}} \sqrt{g_\alpha^h(x)}.$$

Further, if $F(x) = Mx + q$ where M is a $n \times n$ symmetric and positive definite matrix and $q \in \mathbb{R}^n$, we have

$$\|x - x^*\| \leq \sqrt{\frac{1}{\lambda_{\min}(M) - \alpha L}} \sqrt{g_\alpha^h(x)}.$$

The class of functions g_α^h is naturally a subclass of the following family of gap functions given as

$$\hat{g}_\alpha(x) = \sup_{y \in C} \{\langle F(x), x - y\rangle - \alpha\phi(x, y)\}.$$

This class of gap functions was studied by Wu, Florian and Marcotte [18]. In fact the above function can be shown to be a gap function to $VI(F, C)$ if the function ϕ satisfies the following four properties

i) $\phi(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$;
ii) the function ϕ is continuously differentiable on $\mathbb{R}^n \times \mathbb{R}^n$;
iii) for each fixed x the function $y \mapsto \phi(x, y)$ is strongly convex in y in a uniform way in the sense that the modulus of strong convexity is independent of the choice of x;
iv) $\phi(x, y) = 0$ if and only if $y = x$.

In fact using the above four properties of ϕ one can establish that $\hat{g}_\alpha$ is a gap function in a straight forward way and in fact in a more simpler fashion than that of [18] whose proof is motivated by the approach in [10]. The approach that we will just outline depends on the crucial fact that $\nabla_y \phi(x, y) = 0$ if and only if $x = y$. This can be proved very simply based on the property iv). It is in fact trivial to see that $\hat{g}_\alpha(x) \geq 0$ for all $x \in C$. Let $\bar{x} \in C$ be such that $\hat{g}_\alpha(\bar{x}) = 0$. For each $k \in \mathbb{N}$ set

$$y^k = \bar{x} + \frac{1}{k}(y - \bar{x})$$

for any arbitrary (but fixed) $y \in C$. It is also clear that $y_k \in C$ for each k. From the definition of $\hat{g}_\alpha$ we have

$$\langle F(\bar{x}), \bar{x} - y\rangle \leq k\alpha\phi(\bar{x}, y^k). \tag{3}$$

Now by the strong convexity of ϕ in the second variable, we have using the fact that $\phi(\bar{x}, \bar{x}) = 0$

$$-k\alpha\phi(\bar{x}, y^k) \geq \alpha\langle \nabla_y \phi(\bar{x}, y^k), \bar{x} - y\rangle + \frac{\alpha\rho}{k}\|y - \bar{x}\|^2, \tag{4}$$

where $\rho > 0$ is the modulus of strong convexity. Note that since $\phi(\bar{x}, y)$ is strongly convex in y and also differentiable, it is also continuously differentiable and further noting that $\nabla_y \phi(\bar{x}, \bar{x}) = 0$ we see that as $k \to \infty$ from (4) we have

$$\liminf_{k \to \infty} k\alpha\phi(\bar{x}, y^k) \leq 0.$$

Now using (3) we conclude that $\bar{x}$ solves $VI(F, C)$.

For the converse assume that $\bar{x}$ solves $VI(F, C)$ then the fact $\hat{g}_\alpha(\bar{x}) = 0$ follows very simply and we avoid the proof here.

Note that if we set $\phi(x, y) = D_h(y, x)$ where h is a differentiable and strongly convex function then it is simple to see that $D_h(y, x)$ satisfies all the properties of $\phi(x, y)$ listed above. In fact we know by choosing specific forms of h we can get different types of gap functions including the regularized version of Fukushima [10]. An important question that can be asked at this point is as follows. Is the representation of $\phi(x, y)$ as $D_h(y, x)$ is the only possible representation for which all the four properties of ϕ holds true? To the best of our knowledge this remains to be settled.

In fact in order to derive an error bound we have to impose Lipschitz condition on $\nabla_y \phi(x, \cdot)$ for every x and we will also have to assume that the Lipschitz constant does not depend on x. For more details see [18].

It is important to note that in this article we do not discuss the notion of D-gap functions or difference gap functions which is obtained by taking the difference of the same type of regularized gap function but with two different tuning parameters. For more details see Peng [16] and Yamashita, Taji and Fukushima [19]

4. Gap functions for GVI

As compared to the usual $VI(F, C)$ there has been only a few attempts to develop gap functions and error bounds for variational inequalities with multi-valued maps which we term here as $GVI(T, C)$ and $WGVI(T, C)$. One of the important works in this direction is due to Patriksson [15] where he had considered $T(x) = F(x) + \partial f(x)$ for all $x \in \mathbb{R}^n$ where F is a single-valued function and f is a convex function. Patriksson [15] had modified the primal generalized gap function and developed a gap function for $GVI(F + \partial f, C)$. Since $GVI(F + \partial f, C)$ is equivalent to $HVI(F, f, C)$, the gap function introduced in [15] is also a gap function for $HVI(F, f, C)$. Further it is interesting to note that the gap function introduced in [8] for $HVI(F, f, C)$ is different from the one introduced in [15].

Very recently Aussel and Dutta [4] developed gap functions and regularized gap function for GVI and WGVI. However in their work they did not discuss whether the gap functions that they had introduced are well-behaved or not. In their main definitions Aussel and Dutta [4] did not consider the map T to be compact-valued. They defined an auxiliary set-valued map using the set-valued map T in such a way that the auxiliary map is automatically compact-valued. We will only analyze the situation when T is compact-valued. This is primarily motivated by the fact that most important subdifferentials are compact-valued. However it is important to note for example the basic subdifferential or limiting subdifferential (see for example Mordukhovich [13]) is compact-valued when the underlying function is locally Lipschitz but is not convex-valued in general. So we do not assume that T is convex-valued but state this assumption whenever needed. The gap function that we define here for $GVI(T,C)$ is motivated from the gap function introduced by Aussel and Dutta [4] for $WGVI(T,C)$. Let us without loss of generality assume that T is non-empty and compact-valued. The following function was defined in Aussel and Dutta [4].

$$H(x) = \sup_{y \in C} \inf_{x^* \in T(x)} \langle x^*, x - y \rangle.$$

It is proved in [4] that H is a gap function for $WGVI(T,C)$. Further it is mentioned the statement of Proposition 4.1 in [4] that if T is also convex-valued then H is also a gap function for $GVI(T,C)$. It was mentioned in the proof of Proposition 4.1 in [4] that by applying the Sion's minimax theorem (see for example Berge [6]) one can switch the infimum and the supremum and that will lead to the proof that H is a gap function for $GVI(T,C)$. However no details of the proof were provided in [4]. As we will see that working out the details would turn out to be instructive and would lead to a different gap function for $GVI(T,C)$ where we would no longer use the fact that T is convex-valued.

Assuming that T is also convex-valued we will show that H is also a gap function for $GVI(T,C)$. Since $T(x)$ is convex and compact valued by using the Sion's minimax theorem it is immediate that

$$H(x) = \inf_{x^* \in T(x)} \sup_{y \in C} \langle x^*, x - y \rangle.$$

It is simple to see that $H(x) \geq 0$ for all $x \in C$. Further let $\bar{x} \in C$ be such that $H(\bar{x}) = 0$. Then it is clear that

$$0 = \inf_{x^* \in T(\bar{x})} \sup_{y \in C} \langle x^*, \bar{x} - y \rangle.$$

Let us set

$$\beta(x^*, \bar{x}) = \sup_{y \in C} \langle x^*, \bar{x} - y \rangle.$$

Note that as $\bar{x}$ is fixed, $\beta(x^*, \bar{x})$ is a lower-semicontinuous convex function of x^*. Further $\beta(x^*, \bar{x})$ is proper. Note that since $0 = \inf_{x^* \in T(\bar{x})} \beta(x^*, \bar{x})$, for any $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$, there exists $\tilde{x}^* \in T(\bar{x})$ such that $0 \leq \beta(\tilde{x}^*, \bar{x}) < \varepsilon$. Since $T(\bar{x})$ is compact there exists $\hat{x}^*$ such that $\beta(\hat{x}^*, \bar{x}) = 0$. This shows that for all $y \in C$

$$\langle \hat{x}^*, \bar{x} - y \rangle \leq 0.$$

Hence $\bar{x}$ solves $GVI(T,C)$. Conversely, if $\bar{x} \in C$ is such that $\bar{x}$ solves $GVI(T,C)$ then it is clear that there exists $x^* \in T(\bar{x})$ such that

$$\langle x^*, \bar{x} - y \rangle \leq 0 \quad \forall y \in C.$$

This allows us to conclude that $H(\bar{x}) \leq 0$ and thus $H(\bar{x}) = 0$.

Note that in the above proof we do not require the convexity of $T(x)$ once we have switched the infimum with the supremum. This shows that even when T is just compact-valued one can devise a gap function for $GVI(T, C)$. This gap function is given as

$$\hat{H}(x) = \inf_{x^* \in T(x)} \sup_{y \in C} \langle x^*, x - y \rangle.$$

It is however important to note that the compactness of C will be pivotal in guaranteeing the finiteness of H or $\hat{H}$. Thus it is important to have a gap function which remains finite even when C is unbounded. This can be done by regularizing H and $\hat{H}$ in the sense of Fukushima [10]. The regularization for H was done in Aussel and Dutta [4] and for any $\alpha > 0$ the regularized version of H was given as

$$H_\alpha(x) = \sup_{y \in C} \left\{ \inf_{x^* \in T(x)} \langle x^*, x - y \rangle - \frac{\alpha}{2} \|y - x\|^2 \right\}.$$

It was shown in [4] that H_α is finite and is a gap function for $WGVI(T, C)$ and an error bound can also be devised when T is strongly monotone. We refer the reader to [4] for details. It can be shown that H_α is also a lower-semicontinuous function. We will now demonstrate that a regularization of $\hat{H}$ will lead to a finite-valued gap function for $GVI(T, C)$

$$\hat{H}_\alpha(x) = \inf_{x^* \in T(x)} \left\{ \sup_{y \in C} \left\{ \langle x^*, x - y \rangle - \frac{\alpha}{2} \|y - x\|^2 \right\} \right\}.$$

To begin with we will first prove that $\hat{H}_\alpha$ is a gap function.

Theorem 4.1. *Let us consider the problem $GVI(T, C)$ where T is compact-valued. Then for any $\alpha > 0$ the function $\hat{H}_\alpha$ is finite-valued and is a gap function for $GVI(T, C)$.*

Proof. We will first show that $\hat{H}_\alpha$ is finite for any $\alpha > 0$. Observe that we can rewrite $\hat{H}_\alpha$ as $\hat{H}_\alpha(x) = \inf_{x^* \in T(x)} q_\alpha(x^*, x)$ where q_α given below is defined for all pairs $(x^*, x) \in \mathbb{R}^n \times \mathbb{R}^n$.

$$q_\alpha(x^*, x) = \sup_{y \in C} \left\{ \langle x^*, x - y \rangle - \frac{\alpha}{2} \|y - x\|^2 \right\}.$$

Note that q_α can be rewritten as

$$q_\alpha(x^*, x) = - \inf_{y \in C} \left\{ \langle x^*, y - x \rangle + \frac{\alpha}{2} \|y - x\|^2 \right\}.$$

Further it is clear that as (x^*, x) is fixed,

$$y \mapsto \langle x^*, y - x \rangle + \frac{\alpha}{2} \|y - x\|^2$$

is strongly convex and since C is a closed and convex set there exists a unique minimizer $y_\alpha(x^*, x)$ which is given as

$$y_\alpha(x^*, x) = \text{proj}_C \left(x - \frac{1}{\alpha} x^* \right).$$

Hence $q_\alpha(x^*, x)$ is finite for any pair $(x^*, x) \in \mathbb{R}^n \times \mathbb{R}^n$. Further q_α is continuous in (x^*, x) since the projection map is Lipschitz on $\mathbb{R}^n$. Since T is compact-valued this clearly shows that $\hat{H}_\alpha$ is finite-valued.

Our next aim is to show that $\hat{H}_\alpha$ is a gap function. It is clear that $\hat{H}_\alpha(x) \geq 0$ for all $x \in C$. Now let $\bar{x} \in C$ be such that $\hat{H}_\alpha(\bar{x}) = 0$. Consider an arbitrary point $x \in C$ and all points $y(\lambda) = \bar{x} + \lambda(x - \bar{x})$ where $\lambda \in (0, 1)$. Since C is a convex set $y(\lambda) \in C$ for all $\lambda \in (0, 1)$. We have

$$0 = \inf_{x^* \in T(\bar{x})} q_\alpha(x^*, \bar{x}).$$

Since T is compact-valued and we have shown above that q_α is finite and continuous it clear that there exists $w^* \in T(\bar{x})$ such that

$$0 = q_\alpha(w^*, \bar{x}).$$

Thus we have that for all $y \in C$

$$\langle w^*, \bar{x} - y \rangle \leq \frac{\alpha}{2} \|y - x\|^2.$$

In particular for $y = y(\lambda)$ with $\lambda \in (0, 1)$ we have

$$\langle w^*, \bar{x} - x \rangle \leq \frac{\lambda\alpha}{2} \|x - \bar{x}\|^2.$$

As $\lambda \to 0$ we conclude that $\langle w^*, \bar{x} - x \rangle \leq 0$. Since x was arbitrary we see that $\bar{x} \in C$ is a solution of $GVI(T, C)$.

Conversely assume that $\bar{x} \in C$ solves $GVI(T, C)$. Then there exists $p^* \in T(\bar{x})$ such that

$$\langle p^*, \bar{x} - y \rangle \leq 0 \quad \forall y \in C.$$

This shows that $q_\alpha(p^*, \bar{x}) \leq 0$ and thus proving that $\hat{H}_\alpha(\bar{x}) \leq 0$. Hence $\hat{H}_\alpha(\bar{x}) = 0$. This completes the proof. ■

Our next aim is to see if an error bound can be devised for $GVI(T, C)$ in terms of $\hat{H}_\alpha$ when T is μ-strongly monotone. Once that is done then we see whether $\hat{H}_\alpha$ has the desired properties which would make it a suitable choice for devising error bounds.

Theorem 4.2. *Let us consider the problem $GVI(T, C)$, where T is compact-valued and μ-strongly monotone. Let $\alpha > 0$ be such that $2\mu > \alpha$ and let $\bar{x}$ be the unique solution of $GVI(T, C)$. Then for any $x \in C$ we have*

$$\|x - \bar{x}\| \leq \sqrt{\frac{2}{2\mu - \alpha}} \sqrt{\hat{H}_\alpha(x)}.$$

Proof. Observe that for any given $x \in C$ we can find $w^* \in T(x)$ such that

$$q_\alpha(w^*, x) = \hat{H}_\alpha(x).$$

This shows that

$$\hat{H}_\alpha(x) \geq \langle w^*, x - \bar{x} \rangle - \frac{\alpha}{2} \|x - \bar{x}\|^2. \tag{5}$$

Since $\bar{x}$ solves $GVI(T, C)$ there exists $x^* \in T(\bar{x})$ such that

$$\langle x^*, x - \bar{x} \rangle \geq 0. \tag{6}$$

Since T is μ-strongly monotone we have

$$\langle w^*, x - \bar{x} \rangle \geq \langle x^*, x - \bar{x} \rangle + \mu \|x - \bar{x}\|^2.$$

Now using (6) we see that

$$\langle w^*, x - \bar{x} \rangle \geq \mu \|x - \bar{x}\|^2.$$

Combining this with (5) we have

$$\hat{H}_\alpha(x) \geq (\mu - \frac{\alpha}{2}) ||x - \bar{x}||^2.$$

The result follows from the assumption that $2\mu > \alpha$. ■

Now comes the most important question from the computational point of view. Is the gap function $\hat{H}_\alpha$ well-behaved? If we recall our discussion in the beginning of the article then the function $\hat{H}_\alpha$ is well-behaved if for any sequence x_k in C such that $x_k \to \bar{x}$ we must have $\hat{H}_\alpha(x_k) \to 0$. This fact might be not so easy to establish. Observe that

$$\hat{H}_\alpha(x) = \inf_{x^* \in T(x)} q_\alpha(x, x^*),$$

where $q_\alpha(x, x^*)$ is as given in Theorem 4.1. Note that this is a very general form of a parametric optimization problem. Thus it might be quite difficult to guarantee that continuity or at the best the well-behaved-ness of $\hat{H}_\alpha$ with minimal assumptions. However under slightly stringent conditions we can actually show that $\hat{H}_\alpha$ is well behaved. To begin with let us consider the set-valued map $\Lambda : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ which is given as

$$\Lambda(x) = \{x^* \in T(x) : \hat{H}_\alpha(x) = q_\alpha(x, x^*)\}.$$

Note that the assumption that T is non-empty and compact-valued guarantees that Λ is also non-empty and compact-valued. We shall now show that if the set-valued map Λ is lower-semicontinuous in the sense of Berge [6] then one can guarantee that $\hat{H}_\alpha$ is well behaved. Note that by lower-semicontinuity of the mapping Λ we mean that for $x_k \to \bar{x}$ and $\bar{x}^* \in \Lambda(\bar{x})$ there exists $x_k^* \in \Lambda(x_k)$ such that $x_k^* \to \bar{x}^*$. Note that by the definition of $\hat{H}_\alpha$ for any $\bar{x}^* \in \Lambda(\bar{x})$ we have $q_\alpha(\bar{x}, \bar{x}^*) = 0$. Further since if x_k is a sequence in C converging to the solution $\bar{x}$ then by lower-semicontinuity of the set-valued map Λ there exists $x_k^* \in \Lambda(x_k)$ with $x_k^* \to \bar{x}^*$. Thus we have

$$\hat{H}_\alpha(x_k) = q_\alpha(x_k, x_k^*).$$

Now using the continuity of the function q_α we see that

$$\lim_{k\to\infty} \hat{H}_\alpha(x_k) = q_\alpha(\bar{x}, \bar{x}^*) = 0.$$

This shows that $\hat{H}_\alpha$ is well-behaved. Thus we can now summarize our discussion as follows

Theorem 4.3. *Let us consider the problem $GVI(T, C)$, where T is compact-valued. Let $\bar{x}$ be a solution to $GVI(T, C)$ and let $\{x^k\}$ be a sequence in C such that $x^k \to \bar{x}$. Further assume that the set-valued map Λ is lower-semicontinuous in the sense of Berge. Then*

$$\lim_{k\to\infty} \hat{H}_\alpha(x^k) = 0.$$

Hence $\hat{H}_\alpha$ is well-behaved.

Though the assumption of lower-semicontinuity of the set-valued map Λ may appear to be stringent it seems for the present it is difficult to remove this assumption. Thus it remains open to see if there are much simpler conditions on the problem data which will guarantee the function $\hat{H}_\alpha$ to be well behaved.

Acknowledgements. We would like to thank the anonymous referees whose constructive suggestions have vastly improved the presentation of the paper.

References

1. G. Auchmuty, Variational principles for variational inequalities, *Numer. Funct. Anal. Optim.* **10** (1989), no. 9–10, 863–874.
2. A. Auslender, *Optimisation: Methodes Numeriques* (*Maîtrise de Mathématiques et Applications Fondamentales*), Masson, Paris, New York, Barcelona, 1976.
3. A. Auslender, Resolution numerique d'inegalites variationnelles (French), *C. R. Acad. Sci. Paris, Ser. A–B* **276** (1973), 1063–1066.
4. D. Aussel and J. Dutta, On gap functions for multivalued Stampacchia variational inequalities, *J. Optim. Theory Appl.* **149** (2011), no. 3, 513–527.
5. D. Aussel and J. Dutta, Revisiting a gap function for set-valued variational inequalities, preprint 2012.

6. C. Berge, *Topological Spaces: Including a Treatment of Multi-valued Functions, Vector Spaces and Convexity*, Translated from the French original by E. M. Patterson, Reprint of the 1963 translation, Dover Publications, Inc., Mineola, NY, 1997.
7. R. S. Burachik and A. N. Iusem, *Set-Valued Mappings and Enlargements of Monotone Operators*, Springer Optimization and Its Applications, Vol. 8, Springer, New York, 2007.
8. G. Y. Chen, C. J. Goh and X. Q. Yang, On gap functions and duality of variational inequality problems, *J. Math. Anal. Appl.* **214** (1997), no. 2, 658–673.
9. F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer Series in Operations Research, Vol. I, Springer-Verlag, New York, 2003.
10. M. Fukushima, Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems, *Math. Program., Ser. A* **53** (1992), no. 1, 99–110.
11. T. Larsson and M. Patriksson, A class of gap functions for variational inequalities, *Math. Program., Ser. A* **64** (1994), no. 1, 53–79.
12. P. Marcotte, A new algorithm for solving variational inequalities with application to the traffic assignment problem, *Math. Program.* **33** (1985), no. 3, 339–351.
13. B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation*, Springer-Verlag, Berlin, 2006.
14. J. S. Pang, Error bounds in mathematical programming, *Math. Program., Ser. B* **79** (1997), no. 1–3, 299–332.
15. M. Patriksson, Merit functions and descent algorithms for a class of variational inequality problems, *Optimization* **41** (1997), no. 1, 37–55.
16. J. Peng, Equivalence of variational inequality problems to unconstrained minimization, *Math. Program., Ser. A* **78** (1997), no. 3, 347–355.
17. J. P. Vial, Strong and weak convexity of sets and functions, *Math. Oper. Res.* **8** (1983), 231–259.
18. J. H. Wu, M. Florian and P. Marcotte, A general descent framework for the monotone variational inequality problem, *Math. Program., Ser. A* **61** (1993), no. 3, 281–300.
19. N. Yamashita, K. Taji and M. Fukushima, Unconstrained optimization reformulations of variational inequality problems, *J. Optim. Theory Appl.* **92** (1997), no. 3, 439–456.