# Solving Quasi-Variational Inequalities Using the Progressive Decoupling of Linkages

Manoel Jardim* Claudia Sagastizábal† Mikhail Solodov*‡

August 27, 2026

**Abstract**

Inspired by the progressive decoupling of linkages methodology for optimization and variational inequalities, we propose an algorithm for solving quasi-variational inequalities as a sequence of variational inequalities. Our method is shown to converge locally under some regularity conditions and globally when such conditions hold throughout the entire domain. Separately, under other type of assumptions, global convergence with linear rate is also established for the class of quasi-variational inequalities said to have a moving set. Advantageous computational performance is shown for large-scale Walrasian equilibrium problems, a special case of generalized Nash equilibria, as well as for some other instances encountered in related literature.

**Keywords:** quasi-variational inequality; generalized Nash equilibrium problem; progressive decoupling; Spingarn's partial inverse; proximal point method.

## 1 Introduction

The quasi-variational inequality (QVI) is a framework for solving a broad spectrum of variational and equilibrium problems. Given $F : \mathbb{R}^n \to \mathbb{R}^n$ and a set-valued mapping $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, the QVI formulation is as follows:

$$\text{find } \bar{x} \in K(\bar{x}), \text{ such that } \langle F(\bar{x}), x - \bar{x} \rangle \geq 0, \text{ for all } x \in K(\bar{x}), \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ stands for the Euclidean inner product in $\mathbb{R}^n$. The QVI setting is more general than the well-known variational inequalities (VI), the latter corresponding to the case when the set-valued mapping $K(\cdot)$ is constant (the same set $K$ for all $x$).

The first QVI formulation was proposed in [5] and [6] for stochastic impulsive optimal control problems, see also [7]. Since then, the framework has shown itself to be very valuable in economics, for solving generalized Nash equilibrium problems (GNEP) [21], [3], [17], as well as in some engineering and transportation applications, see [28] and [38] respectively. The monograph [19] discusses the mathematical theory of QVIs and VIs in detail.

Much of the literature on QVIs concerns theoretical results, such as the existence and/or uniqueness of solutions. In the context of QVIs in Banach spaces, several existence results have been established for specific problems, in a way similar to developments in partial differential equations theory, often together with uniqueness and regularity of solutions. Some examples can be found in [4], [22], and [29]. An existence theorem for a general QVI in Banach spaces is presented in [26]. General theorems establishing existence of solutions for QVIs in Euclidean spaces can be found in [8] and [2]. There are also existence results for QVIs in specific applications, see for example [3] and [11]. However, there are only a few specific theorems addressing uniqueness, such as the results in [33] and [14].

*IMPA, Estrada Dona Castorina 110, Rio de Janeiro, RJ 22460-320, Brazil. manoel.jardim@impa.br
†IMECC, Unicamp, Rua Sérgio Buarque de Holanda, Campinas, SP 13083-859, Brazil. sagastiz@unicamp.br
‡solodov@impa.br

More recently, in order to address large-scale structured problems, the QVI Dantzig-Wolfe decomposition in [23] demonstrated promising computational efficiency and scalability. Nevertheless, in its so-called master subproblem each Dantzig-Wolfe iteration still solves a QVI, albeit a simpler one than the original one (the second subproblem to be solved at each Dantzig-Wolfe iteration is a VI, often separable).

Our goal in this work is to design an algorithm capable of handling large-scale instances of (1.1) through the iterative solution of VI problems. To this end, we employ a decomposition approach different from the Dantzig-Wolfe paradigm, inspired by the progressive decoupling of linkages proposed in [37] for optimization and VI problems. After lifting (1.1) into a higher dimensional space, a common technique to induce decomposition, the linkage is given by a subspace constraint. In each iteration, progressive decoupling consists of first solving a proximal subproblem in the larger space, without the aforementioned constraint, and then performing a projection onto the subspace to recover feasibility of the linkage. To extend this mechanism to our setting, we first rewrite the QVI in a favorable format. Thanks to this reformulation, the subproblem solved by the progressive decoupling approach is just one VI, as desired. The subsequent projection onto the subspace has an easy explicit expression.

We expect our new proposal to outperform the Dantzig-Wolfe approach [23] when the master QVI subproblem therein is more difficult to solve, which occurs when the problem has more constraints and QVI subproblems are not too small. The expectation is confirmed by the numerical experiments in Section 7, where it is also shown that the situation is reversed for instances with fewer constraints. For computational assessment, we solve Walrasian equilibrium problems formulated as GNEPs, using their optimality conditions that give rise to a QVI, as in [21], [18], [16]. We compare our method with a direct application of the PATH solver in GAMS [20] and with the Dantzig-Wolfe approach of [23]. Additionally, we solve several other QVI examples from [16] and [9] that are useful to analyze some convergence features of our method, including sensitivity with respect to the subproblems' proximal parameter.

When dealing with proximal subproblems, one difficulty which is inherent in QVIs is the lack of monotonicity-type properties of the associated mapping. We briefly discuss this next. Let $K(x)$ be convex for each $x$. For a convex set $D$, the normal cone to $D$ at $x$ is the set $N_D(x) = \{v \in \mathbb{R}^n : \langle v, y - x\rangle \leq 0, \text{for all } y \in D\}$ when $x \in D$, and $N_D(x) = \emptyset$ otherwise. Then the QVI (1.1) is equivalent to finding $\bar{x}$ solving the generalized equation

$$0 \in F(\bar{x}) + N_{K(\bar{x})}(\bar{x}). \tag{1.2}$$

Because $K$ in a QVI is not a constant set, but varies with $x$, the mapping $N_{K(\cdot)}(\cdot) : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ may lack properties suitable for a direct application of proximal point and related techniques, such as the progressive decoupling method. For this reason, special attention is required. Following [35] we refer to such a mapping as the putative normal cone. In particular, this mapping is not monotone, even in very simple examples. Moreover, even weaker properties like local monotonicity around a solution or elicited monotonicity, do not appear to be amenable to the QVI case (a further discussion is given in Section 4; see in particular Example 4.1 therein). For this reason, part of our analysis relies on local results. We shall prove local convergence, under suitable regularity conditions and global convergence when the conditions hold throughout the entire domain. This analysis invokes variants of the progressive decoupling method proposed in [15]. Note that even though the method proposed in [37] allows elicitation of monotonicity, the progressive decoupling+ from [15] covers more general cases and provides both local and global convergence results under a calmness condition. In addition, we establish global convergence with linear rate for a specific instance of (1.1), namely the moving set case, by direct analysis not involving the theory of [15].

The rest of the paper is organized as follows. In Section 2 and Section 3, we develop and formulate the algorithm. In Section 4, we introduce the regularity conditions necessary for the convergence analysis. Section 5 and Section 6 provide global and local convergence analysis, respectively. The final Section 7 is devoted to the numerical assessment.

We introduce the following notation. Denoting the Euclidean norm by $\|\cdot\|$, the closed unit ball in $\mathbb{R}^n$ is $\mathbb{B}^n = \{v \in \mathbb{R}^n : \|v\| \leq 1\}$. For a vector subspace $\mathbf{S} \subset \mathbb{R}^n$, its orthogonal subspace is $\mathbf{S}^\perp = \{u \in \mathbb{R}^n : \langle u, v\rangle = 0, \text{for all } v \in \mathbf{S}\}$. It then holds that $\mathbb{R}^n = \mathbf{S} \oplus \mathbf{S}^\perp = \{v \in \mathbb{R}^n : v = v_\mathbf{S} + v_{\mathbf{S}^\perp}, v_\mathbf{S} \in \mathbf{S}, v_{\mathbf{S}^\perp} \in \mathbf{S}^\perp\}$ is the direct sum decomposition of $\mathbb{R}^n$, where

$v_{\mathbf{S}} = \mathbf{P}_{\mathbf{S}}(v)$ and $v_{\mathbf{S}^\perp} = \mathbf{P}_{\mathbf{S}^\perp}(v)$ are the projections of $v$ onto $\mathbf{S}$ and $\mathbf{S}^\perp$, respectively.

For a set-valued mapping $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, its domain is given by $\operatorname{dom}\Phi = \{x \in \mathbb{R}^n : \Phi(x) \neq \emptyset\}$, its graph is $\operatorname{gph}\Phi = \{(x,u) : x \in \operatorname{dom}\Phi, u \in \Phi(x)\}$, its range is $\operatorname{rge}(\Phi) = \{u \in \mathbb{R}^n : u \in \Phi(x) \text{ for some } x \in \operatorname{dom}\Phi\}$, and its set of fixed points is given by $\operatorname{Fix}(\Phi) = \{x : x \in \Phi(x)\}$. The mapping $\Phi$ is $\sigma$-strongly monotone if there exists $\sigma > 0$ such that $\langle v - u, y - x\rangle \geq \sigma\|y - x\|^2$, for all $(y,v),(x,u) \in \operatorname{gph}\Phi$. The mapping $\Phi$ is monotone if the inequality holds with $\sigma = 0$. A monotone set-valued mapping is maximally monotone if its graph is not properly contained in the graph of any other monotone mapping.

A set-valued mapping $T : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is $e$-elicitable monotone with respect to the subspace $\mathbf{S}$ if $T + eP_{\mathbf{S}^\perp}$ is maximally monotone for some $e \geq 0$. Given a proximal parameter $r > 0$, the resolvent of $T$ with parameter $r$ is defined as $\mathcal{J}_{\frac{1}{r}T} := (I + \frac{1}{r}T)^{-1} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, a possibly set-valued mapping if $T$ is non-monotone.

A set-valued mapping $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is inner semicontinuous at $\bar{u} \in \operatorname{dom} K$ if for all $u^k \to \bar{u}$ and $\bar{v} \in K(\bar{u})$, there exists a sequence $v^k \to \bar{v}$ with $v^k \in K(u^k)$. In our QVI 1.1, the set-valued mapping $K(\cdot)$ is assumed to be closed and convex-valued, meaning that the sets $K(x)$ are closed and convex for each $x \in \operatorname{dom} K$. Moreover, $K(\cdot)$ is assumed to be inner semicontinuous and to have a closed graph. Finally, the variational inequality obtained when $K$ in (1.1) is a fixed set is denoted by $\operatorname{VI}(F, K)$.

# 2 QVI Progressive Decoupling

A useful mechanism to induce decomposition in a problem is to make copies of variables appearing in coupling constraints, in a manner that separable structures can be exploited. The mechanism is defined in a decision space of larger dimensions, but in compensation each iteration yields separate subproblems, generally simpler to solve, possibly in parallel. Decisions in the original space are obtained by forcing the copies to be equal, and the iterative process continues.

## 2.1 The considered setting

We denote with bold font all vectors, mappings, and sets in $\mathbb{R}^{2n}$, the space of copies. We introduce the subspace $\mathbf{S} = \left\{\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R}^n : x_1 = x_2\right\}$, and its orthogonal complement $\mathbf{S}^\perp = \left\{\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \in \mathbb{R}^n \times \mathbb{R}^n : y_1 = -y_2\right\}$. The associated projection operators are

$$\mathbf{P}_{\mathbf{S}}(\mathbf{x}) = \frac{1}{2}\begin{bmatrix} x_1 + x_2 \\ x_1 + x_2 \end{bmatrix} \text{ and } \mathbf{P}_{\mathbf{S}^\perp}(\mathbf{x}) = \frac{1}{2}\begin{bmatrix} x_1 - x_2 \\ x_2 - x_1 \end{bmatrix} \text{ for any } \mathbf{x} \in \mathbb{R}^{2n}. \tag{2.1}$$

The QVI formulation (1.2) in the space of copies, is given by

$$(\mathbf{x}, \mathbf{y}) \in \mathbf{S} \times \mathbf{S}^\perp, \text{ with } \quad \mathbf{y} \in \mathbf{T}(\mathbf{x}) = (F(x_1) + N_{K(x_2)}(x_1)) \times \{0\}, \tag{2.2}$$
$$\text{and} \qquad x_1 \in K(x_2). \tag{2.3}$$

This is Spingarn's partial inverse problem, where the goal is to find a pair in the graph of the operator such that the point and its image respectively lie in a subspace and its orthogonal complement. The transformation was proposed in [39] for convex separable programming problems, and the progressive decoupling method for VIs [37] is based on this type of construction.

To analyze the framework in QVI setting, we begin by examining one iteration of the progressive decoupling of linkages when applied to (2.2)–(2.3). At iteration $k$, given $(\mathbf{x}^k, \mathbf{y}^k) \in \mathbf{S} \times \mathbf{S}^\perp$ and a proximal parameter $r_k > 0$, the next iterate $(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}) \in \mathbf{S} \times \mathbf{S}^\perp$ is generated in two steps, by solving a proximal subproblem and projecting. Specifically,

$$\text{first,} \quad \hat{\mathbf{x}}^k \text{ is a solution to} \quad \mathbf{0} \in \mathbf{T}(\mathbf{x}) - \mathbf{y}^k + r_k(\mathbf{x} - \mathbf{x}^k), \tag{2.4}$$
$$\text{then,} \quad \text{(2.1) gives} \quad \mathbf{x}^{k+1} = \mathbf{P}_{\mathbf{S}}(\hat{\mathbf{x}}^k) \text{ and } \mathbf{y}^{k+1} = \mathbf{y}^k - r_k\mathbf{P}_{\mathbf{S}^\perp}(\hat{\mathbf{x}}^k). \tag{2.5}$$

Although the original progressive decoupling method is formulated for a fixed proximal parameter, we write the iteration with a variable parameter for greater generality. We restrict to the case of a constant parameter whenever required for the convergence analysis, except for the global convergence result in the moving-set setting (where we show that variable parameters are admissible indeed).

Two important comments are now in order. To begin with, because a solution $\hat{\mathbf{x}}^k$ to (2.4) satisfies the inclusion $\mathbf{0} \in \mathbf{T}(\hat{\mathbf{x}}^k) - \mathbf{y}^k + r_k(\hat{\mathbf{x}}^k - \mathbf{x}^k)$, it holds that

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in \begin{bmatrix} F(\hat{x}_1^k) + r_k\hat{x}_1^k - y_1^k - r_k x_1^k + N_{K(\hat{x}_2^k)}(\hat{x}_1^k) \\ r_k\hat{x}_2^k - y_2^k - r_k x_2^k \end{bmatrix}.$$

Thus, at iteration $k$, the two components $\hat{x}_1^k$ and $\hat{x}_2^k$ in (2.4) are such that $\hat{x}_2^k = x_2^k + \frac{1}{r_k}y_2^k$, while $\hat{x}_1^k$ solves VI $(F + r_kI - y_1^k - r_kx_1^k, K^k)$, where we defined the set $K^k = K(\hat{x}_2^k)$.

Second, by the projection formula in (2.5), the next iterates are

$$\mathbf{x}^{k+1} = \frac{1}{2}\begin{bmatrix} \hat{x}_1^k + \hat{x}_2^k \\ \hat{x}_1^k + \hat{x}_2^k \end{bmatrix} \implies x_1^{k+1} = x_2^{k+1} = \frac{1}{2}(\hat{x}_1^k + x_2^k + \frac{1}{r_k}y_2^k),$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k - \frac{r_k}{2}\begin{bmatrix} \hat{x}_1^k - \hat{x}_2^k \\ \hat{x}_2^k - \hat{x}_1^k \end{bmatrix} \implies \begin{array}{l} y_1^{k+1} = y_1^k - \frac{r_k}{2}(\hat{x}_1^k - x_2^k - \frac{1}{r_k}y_2^k), \\ y_2^{k+1} = y_2^k - \frac{r_k}{2}(x_2^k + \frac{1}{r_k}y_2^k - \hat{x}_1^k). \end{array}$$

Since $\mathbf{x}^k \in \mathbf{S}$ and $\mathbf{y}^k \in \mathbf{S}^\perp$, the respective components satisfy $x_1^k = x_2^k$ and $y_2^k = -y_1^k$. Thus, renaming $x_1^k = x_2^k = x^k$, and $y_1^k = y^k, y_2^k = -y^k$ yields the expression $\hat{x}_2^k = x^k - \frac{1}{r_k}y^k$. We can also rename $\hat{x}_1^k = \hat{x}^k$.

Altogether, the proximal subproblem of the iteration consists in finding

$$\hat{x}^k \text{ solving the VI}\Big(F + r_kI - y^k - r_kx^k, K(x^k - \frac{1}{r_k}y^k)\Big).$$

## 2.2 A first algorithmic setting

Since we have characterized $\hat{x}^k$ as a solution to a VI with operator $F + r_kI - y^k - r_kx^k$ and set $K^k = K(x^k - \frac{1}{r_k}y^k)$, the subproblems are well defined without further assumptions/considerations, at least locally. To see this, note that as $K$ is inner semicontinuous in $\bar{x}$, then $\bar{x} \in \operatorname{int}\operatorname{dom} K$. Therefore, $K^k$ will be non-empty whenever $(x^k, y^k)$ is sufficiently close to $(\bar{x}, 0)$. If, in addition, $F$ were monotone, the VI mapping $F(\cdot) + r_kI - y^k - r_kx^k$ would be strongly monotone. In this case, by Proposition 2.3.3-(b) in [19], the corresponding $k$th VI subproblem has the unique solution $\hat{x}^k$.

---

**Algorithm 1** Progressive decoupling for QVI 1.1

---

**Require:** $x^0, y^0 \in \mathbb{R}^n, \text{tol} \geq 0, \varepsilon^0 > \text{tol}$.
**Ensure:** Convergence to solutions of the QVI.
  Set $k \leftarrow 0$.
  **while** $\varepsilon^k > \text{tol}$ **do**
    Choose $r_k > 0$ and set $K^k = K(x^k - \frac{1}{r_k}y^k)$;
    find $\hat{x}^k$ as a solution to VI $(F + r_kI - y^k - r_kx^k, K^k)$;
    set $x^{k+1} = \frac{1}{2}(\hat{x}^k + x^k - \frac{1}{r_k}y^k)$;
    set $y^{k+1} = y^k - \frac{r_k}{2}(\hat{x}^k - x^k + \frac{1}{r_k}y^k)$.
    STOPPING CRITERION: Compute $\varepsilon^k = \|x^{k+1} - x^k\| + \frac{1}{r_k}\|y^{k+1} - y^k\|$;
    UPDATE: $k \leftarrow k + 1$;
  **end while**

---

Note that the iterative procedure given in Algorithm 1 solves one variational inequality in $\mathbb{R}^n$ per iteration, the remaining calculations being simple algebraic operations. The stopping criterion is based on the convergence of the corresponding sequences, whose analysis is provided in Sections 5 and 6. For now, we provide some insights as to why, if the method converges in some sense, then a solution to (1.1) is obtained (of course, approximate solutions for a positive tolerance $\text{tol} > 0$).

**Proposition 2.1** (Meaning of the stopping criterion)**.** *Suppose that in QVI 1.1 the operator $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous and $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is an inner semicontinuous set-valued mapping with closed graph.*

*Suppose, in addition, the proximal parameters in Algorithm 1 are chosen so that $0 < \check{r} \leq r_k \leq \hat{r} < \infty$ for all $k$.*

*If the process generates a sequence $\{y^k\} \to 0$, then every accumulation point $\bar{x}$ of the sequence $\{x^k\}$ is a solution of the QVI (1.1). Moreover, whenever $\{x^k\} \to \bar{x}$ the sequence $\{y^k\} \to 0$, and $\bar{x}$ solves the QVI (1.1).*

*Proof.* Suppose that $\{y^k\} \to 0$ and consider $k \to \infty$. As $y^{k+1} = -\frac{r_k}{2}(\hat{x}^k - x^k - \frac{1}{r_k}y^k)$, our conditions on $r_k$ imply that $\hat{x}^k - x^k \to 0$.

For $\bar{x}$ an accumulation point of $\{x^k\}$, consider the subsequence $\{x^{k_j}\} \to \bar{x}$ as $j \to \infty$. Then it also holds that $\{\hat{x}^{k_j}\} \to \bar{x}$ as $j \to \infty$. Since for all $k$ the points $\hat{x}^k \in K(x^k - \frac{1}{r_k}y^k)$ and $\operatorname{gph} K$ is closed, passing to the limit along the subsequence $\{k_j\}$ as $j \to \infty$, we conclude that $\bar{x} \in K(\bar{x})$.

Next, take an arbitrary $u \in K(\bar{x})$. By the inner semicontinuity of $K(\cdot)$, together with the relation $x^{k_j} - \frac{1}{r_{k_j}}y^{k_j} \to \bar{x}$, there exists some subsequence $\{u^{k_j}\} \to u$ such that $u^{k_j} \in K^{k_j} = K(x^{k_j} - \frac{1}{r_{k_j}}y^{k_j})$ for all $j$. Since $\hat{x}^{k_j}$ solves the corresponding VI, it holds that

$$\left\langle F(\hat{x}^{k_j}) + r_{k_j}\hat{x}^{k_j} - y^{k_j} - r_{k_j}x^{k_j}, u^{k_j} - \hat{x}^{k_j} \right\rangle \geq 0.$$

Passing onto the limit as $j \to \infty$ yields that $\langle F(\bar{x}), u - \bar{x}\rangle \geq 0$, a relation that holds for any $u \in K(\bar{x})$. Hence, $\bar{x}$ solves the QVI (1.1), as claimed.

Finally, suppose that $x^k \to \bar{x}$. The result follows, because $x^{k+1} - x^k \to 0$ and

$$y^{k+1} = -\frac{r_k}{2}(\hat{x}^k - x^k - \frac{1}{r_k}y^k) = -\frac{r_k}{2}(\hat{x}^k + x^k - \frac{1}{r_k}y^k - 2x^k) = -r_k(x^{k+1} - x^k) \to 0.$$

That this relation implies that $\bar{x}$ is a solution to (1.1) was established above. ☐

At this stage, it is convenient to recall the relation between the proximal subproblem in Algorithm 1, written in the original space ($\mathbb{R}^n$), and its equivalent expression involving the resolvent of the operator in (2.2), defined in the lifted space of copies, that is:

$$\hat{x}^k \text{ solves the VI} (F + r_k I - y^k - r_k x^k, K^k) \iff \hat{\mathbf{x}}^k = \begin{bmatrix} \hat{x}^k \\ x^k - \frac{1}{r_k}y^k \end{bmatrix} \in \mathcal{J}_{\frac{1}{r_k}\mathbf{T}}\Big(\mathbf{x}^k + \frac{1}{r_k}\mathbf{y}^k\Big).$$

Together with (2.1), we see that, when written in the lifted space of copies, the $k$th iteration of Algorithm1 performs the following steps

$$\begin{cases} \hat{\mathbf{x}}^k \in \mathcal{J}_{\frac{1}{r_k}\mathbf{T}}(\mathbf{x}^k + \frac{1}{r_k}\mathbf{y}^k) \\ \mathbf{x}^{k+1} = \mathbf{P}_{\mathbf{S}}(\hat{\mathbf{x}}^k) \\ \mathbf{y}^{k+1} = \mathbf{y}^k - r_k \mathbf{P}_{\mathbf{S}^\perp}(\hat{\mathbf{x}}^k)\,. \end{cases} \tag{2.6}$$

This rewriting will be useful in the following section to introduce a second algorithmic approach for solving (1.1), when written in its equivalent form (2.2)-(2.3).

# 3 QVI Progressive Decoupling+

The progressive decoupling method [37] handles non-monotonicity of the mapping $\mathbf{T}$ through elicitation within the linkage subspace. Since that setting does not accommodate the lack of monotonicity of the putative cone (see Section 4 and Example 4.1), we consider instead the weaker concept of semimonotonicity, introduced in [15].

The modified variant of the progressive decoupling approach presented in [15] can be employed when elicitation is not applicable. The resulting method, called progressive decoupling+, incorporates in the original framework separate relaxation parameters, $\lambda_x$ and $\lambda_y$ when updating the iterates in (2.6). As long as the proximal parameter is fixed ($r_k \equiv r$ for all iterations), the mechanism allows for a controlled form of non-monotonicity along the subspace directions and their complements.

The approach in [15] applied to our QVI problem writes down as follows. Given $(\mathbf{x}^k, \mathbf{y}^k) \in \mathbf{S} \times \mathbf{S}^\perp$, a proximal parameter $r > 0$, and relaxation parameters $\lambda_x, \lambda_y$, the iteration of the progressive decoupling+ defines

$$\begin{cases} \hat{\mathbf{x}}^k \in \mathcal{J}_{\frac{1}{r}\mathbf{T}}(\mathbf{x}^k + \frac{1}{r}\mathbf{y}^k) \\ \mathbf{x}^{k+1} = (1-\lambda_x)\mathbf{x}^k + \lambda_x \mathbf{P}_{\mathbf{S}}(\hat{\mathbf{x}}^k) \\ \mathbf{y}^{k+1} = \mathbf{y}^k - \lambda_y r \mathbf{P}_{\mathbf{S}^\perp}(\hat{\mathbf{x}}^k). \end{cases}$$

We see that taking $\lambda_x = \lambda_y = 1$ gives the progressive decoupling steps of Algorithm 1, written in the form (2.6) and fixing $r_k = r$ (in the progressive decoupling variant with elicitation the respective values of the parameters are $\lambda_x = 1$ and $\lambda_y = 1 + \frac{e}{r}$).

Algorithm 2 stated below results from [15], after lifting (1.1) to the space of copies as in Section 2.1, using the operator $\mathbf{T}$ given in (2.2).

---

**Algorithm 2** Progressive decoupling+ for QVI 1.1

---

**Require:** $x^0, y^0 \in \mathbb{R}^n, r > 0, \lambda_x > 0, \lambda_y > 0, \mathrm{tol} \geq 0, \varepsilon^0 > \mathrm{tol}$.
**Ensure:** Convergence to solutions of the QVI.
Set $k \leftarrow 0$.
**while** $\varepsilon^k > \mathrm{tol}$ **do**
  Set $K^k = K(x^k - \frac{1}{r}y^k)$;
  find $\hat{x}^k$ as a solution to VI $(F + rI - y^k - rx^k, K^k)$;
  set $x^{k+1} = (1-\lambda_x)x^k + \lambda_x \left(\frac{1}{2}(\hat{x}^k + x^k - \frac{1}{r}y^k)\right)$;
  set $y^{k+1} = y^k - \lambda_y \left(\frac{r}{2}(\hat{x}^k - x^k + \frac{1}{r}y^k)\right)$.
  STOPPING CRITERION: Compute $\varepsilon^k = \|x^{k+1} - x^k\| + \frac{1}{r}\|y^{k+1} - y^k\|$;
  UPDATE: $k \leftarrow k + 1$;
**end while**

---

For ensuring convergence, the parameters $r$ and $\lambda_x, \lambda_y$ have to satisfy certain conditions, which will appear in due course, after the study of semimonotonicity in our QVI setting in the next section.

# 4 On semimonotonicity in QVI setting

Since the progressive decoupling of linkages is an application of the proximal point method (PPM) [37], its convergence properties follow from those of that method under appropriate assumptions. Convergence of PPM depends on monotonicity-related properties (perhaps local monotonicity, perhaps elicitation), which are not easily achievable in the presence of putative normal cones.

## 4.1 Regularity Assumptions

The structure of the mapping $\mathbf{T}$ in (2.2) is not amenable to monotonicity, but the assumptions required by the progressive decoupling+ are less demanding and turn out to be reasonable for quasi-variational inequalities.

We illustrate the situation with the following simple example, in which we have neither monotonicity –nor local monotonicity around the solution –nor elicitation for any parameter. However, as will be shown later on in Example 5.1, we have semimonotonicity at the solution point in the sense of [15], stated in Definition 4.2.

**Example 4.1.** Consider QVI (1.1) defined with $n = 2$, the identity operator $F(x) = x$ and the set-valued mapping $K(x) = \{tx : t \in \mathbb{R}\} \subset \mathbb{R}^2$.

It is easy to see that

$$N_{K(x)}(x) = [x]^\perp := \{h \in \mathbb{R}^2 : \langle h, x \rangle = 0\}.$$

Then the unique solution to QVI (1.2), i.e., $0 \in F(x) + N_{K(x)}(x)$, equivalent to (1.1), is $\bar{x} = 0$.

To derive the operator in the lifted space, first note that given $x_1, x_2 \in \mathbb{R}^2$,

$$N_{K(x_2)}(x_1) = \begin{cases} [x_2]^\perp := \{h \in \mathbb{R}^2 : \langle h, x_2\rangle = 0\}, \text{ when } x_1 \in K(x_2), \\ \emptyset, \text{ when } x_1 \notin K(x_2). \end{cases}$$

It follows that the mapping $\mathbf{T}$ from (2.2) is given by

$$\mathbf{T} : \mathbb{R}^4 \rightrightarrows \mathbb{R}^4, \mathbf{T}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{cases} (x_1 + [x_2]^\perp) \times \{0\}, \text{ if } x_1 \text{ and } x_2 \text{ are parallel}, \\ \emptyset, \text{ otherwise}. \end{cases}$$

Furthermore, keeping (2.1) in mind, for any $\varepsilon > 0, \delta > 0$,

$$\mathbf{T}\left(\begin{bmatrix} \varepsilon \\ 0 \\ -\varepsilon \\ 0 \end{bmatrix}\right) = \left(\begin{bmatrix} \varepsilon \\ 0 \end{bmatrix} + \left\{\begin{bmatrix} 0 \\ t \end{bmatrix} \in \mathbb{R}^2, t \in \mathbb{R}\right\}\right) \times \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\} = \left\{\begin{bmatrix} \varepsilon \\ t \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^4, t \in \mathbb{R}\right\},$$

while, similarly,

$$\mathbf{T}\left(\begin{bmatrix} 0 \\ \delta \\ 0 \\ -\delta \end{bmatrix}\right) = \left\{\begin{bmatrix} t \\ \delta \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^4 : t \in \mathbb{R}\right\}.$$

Therefore,

$$\mathbf{P}_{\mathbf{S}^\perp}\left(\begin{bmatrix} \varepsilon \\ 0 \\ -\varepsilon \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \varepsilon \\ 0 \\ -\varepsilon \\ 0 \end{bmatrix}, \quad \mathbf{P}_{\mathbf{S}^\perp}\left(\begin{bmatrix} 0 \\ \delta \\ 0 \\ -\delta \end{bmatrix}\right) = \begin{bmatrix} 0 \\ \delta \\ 0 \\ -\delta \end{bmatrix},$$

and, hence,

$$\left(\begin{bmatrix} 0 \\ \delta \\ 0 \\ -\delta \end{bmatrix}, \begin{bmatrix} t_2 \\ \delta \\ 0 \\ 0 \end{bmatrix} + e\begin{bmatrix} 0 \\ \delta \\ 0 \\ -\delta \end{bmatrix}\right), \left(\begin{bmatrix} \varepsilon \\ 0 \\ -\varepsilon \\ 0 \end{bmatrix}, \begin{bmatrix} \varepsilon \\ t_1 \\ 0 \\ 0 \end{bmatrix} + e\begin{bmatrix} \varepsilon \\ 0 \\ -\varepsilon \\ 0 \end{bmatrix}\right) \in \operatorname{gph}(\mathbf{T} + e\mathbf{P}_{\mathbf{S}^\perp}), \text{ for } t_1, t_2 \in \mathbb{R}.$$

In particular, given any elicitation parameter $e > 0$,

$$\left\langle \begin{bmatrix} t_2 \\ (1+e)\delta \\ 0 \\ -e\delta \end{bmatrix} - \begin{bmatrix} (1+e)\varepsilon \\ t_1 \\ -\varepsilon e \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \delta \\ 0 \\ -\delta \end{bmatrix} - \begin{bmatrix} \varepsilon \\ 0 \\ -\varepsilon \\ 0 \end{bmatrix} \right\rangle = (1+2e)(\delta^2 + \varepsilon^2) - t_1\delta - t_2\varepsilon < 0, \tag{4.1}$$

$$\text{whenever } t_1\delta + t_2\varepsilon > (1+2e)(\delta^2 + \varepsilon^2).$$

We conclude that for this example, $\mathbf{T} + e\mathbf{P}_{\mathbf{S}^\perp}$ cannot be monotone for any $e > 0$. Thus, $\mathbf{T}$ is not an $e$-elicitable mapping and, moreover, taking $e = 0$, shows that $\mathbf{T}$ is not monotone either. Monotonicity cannot be ensured even locally around the solution $(0, 0) \in \mathbb{R}^4 \times \mathbb{R}^4$, because we can choose $\delta, \varepsilon$ above as small as we wish, without altering the conclusion. On the other hand, the semimonotonicity condition required by progressive decoupling+ is readily satisfied for this example, as will be illustrated in Example 5.1 below. □

As the mapping $\mathbf{T} : \mathbb{R}^{2n} \rightrightarrows \mathbb{R}^{2n}$ cannot be expected to be maximally monotone, locally monotone, or elicitable in the sense of [37], the results on convergence of PPM cannot be applied to progressive decoupling in our setting. Instead, we shall employ the framework of [15], which is based on the notion of semimonotonicity.

**Definition 4.2** (Semimonotonicity)**.** Given a mapping $\Phi : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$, a subspace $\mathcal{V} \subset \mathbb{R}^m$ and scalars $\mu^\perp, \mu > 0$, the mapping $\Phi$ is $(\mu^\perp, \mu)$- semimonotone at $(x^*, y^*) \in \operatorname{gph}\Phi$ on an open set $\mathcal{U} \ni (x^*, y^*)$ if, for all $(x, y) \in \operatorname{gph}\Phi \cap \mathcal{U}$ the following holds:

$$\langle y - y^*, x - x^*\rangle + \mu^\perp \|\mathbf{P}_{\mathcal{V}^\perp}(x) - \mathbf{P}_{\mathcal{V}^\perp}(x^*)\|^2 + \mu\|\mathbf{P}_{\mathcal{V}}(y) - \mathbf{P}_{\mathcal{V}}(y^*)\|^2 \geq 0\,.$$

For our mapping $\mathbf{T}$ in (2.2) and the subspace $\mathbf{S}$, this property will be satisfied at $(\bar{\mathbf{x}}, \mathbf{0})$ if there exist $\delta_x, \delta_y > 0$ such that

$$\langle y, x_1 - \bar{x}\rangle + \frac{\mu^{\perp}}{2}\|x_1 - x_2\|^2 + \frac{\mu}{2}\|y\|^2 \geq 0 \quad \begin{array}{l}\text{for all } x_1, x_2 \in \bar{x} + \delta_x \mathbb{B}^n \text{ and } y \in \delta_y \mathbb{B}^n \\ \text{such that } y \in F(x_1) + N_{K(x_2)}(x_1).\end{array} \tag{4.2}$$

Recalling that Algorithm 1 is a particular case of Algorithm 2, to show convergence for both of our QVI progressive decoupling methods, we next state conditions ensuring that the property (4.2) of semimonotonicity holds.

To this end, given parameters $p, q \in \mathbb{R}^n$, we introduce the following parametric VI and QVI solution mappings:

$$\begin{array}{lcl} x \in \mathbf{M_{vi}}(p, q) & \iff & 0 \in p + F(x) + N_{K(q)}(x), \\ x \in \mathbf{M_{qvi}}(p, q) & \iff & 0 \in p + F(x) + N_{K(x+q)}(x). \end{array} \tag{4.3}$$

Stability properties of parameterized VIs and QVIs have been studied in [36], [31], [34], [32]. Some results are based on metric subregularity theory, which is equivalent to a calmness property for the solution mapping.

We now introduce the regularity conditions for our mappings. For the convergence analysis, it suffices that either of the two assumptions below holds. The difference between the local and global analyses is that the former requires subregularity only locally, while the latter requires the property to hold on the entire domain.

**Assumption 4.3** ($\mathbf{M_{vi}}$-calmness)**.** *Let $\bar{x}$ be a solution of QVI* (1.1)*, meaning that $\bar{x} \in \mathbf{M_{vi}}(0, \bar{x})$. The mapping $\mathbf{M_{vi}}$ is locally single-valued and calm at $((0, \bar{x}), \bar{x})$: there exist neighborhoods $V_p \ni 0, V_q \ni \bar{x}$ and constants $L_p, L_q \geq 0$ with $L_q < 1/2$ such that $\mathbf{M_{vi}}$ is single-valued in $V_p \times V_q$ and*

$$\|\mathbf{M_{vi}}(p, q) - \mathbf{M_{vi}}(0, \bar{x})\| \leq L_p\|p\| + L_q\|q - \bar{x}\|, \text{ for all } (p, q) \in V_p \times V_q.$$

**Assumption 4.4** ($\mathbf{M_{qvi}}$-calmness)**.** *Let $\bar{x}$ be a solution of QVI* (1.1)*, meaning that $\bar{x} \in \mathbf{M_{qvi}}(0, 0)$. The mapping $\mathbf{M_{qvi}}$ is locally single-valued and calm at $((0, 0), \bar{x})$: there exist neighborhoods $V_p \ni 0, V_q \ni 0$ and constants $L_p, L_q \geq 0$ with $L_q < 1$ such that $\mathbf{M_{qvi}}$ is single-valued in $V_p \times V_q$ and*

$$\|\mathbf{M_{qvi}}(p, q) - \mathbf{M_{qvi}}(0, 0)\| \leq L_p\|p\| + L_q\|q\|, \text{ for all } (p, q) \in V_p \times V_q.$$

*Remark* 4.5.

(i) When $F$ is strongly monotone, the mapping $\mathbf{M_{vi}}$ is always single-valued, since uniqueness of the solution of the VI is ensured by Theorem 2.3.3-(b) in [19].

(ii) Assumption 4.3 does not require uniqueness or local uniqueness of solutions to QVI (1.1). Indeed, $\mathbf{M_{vi}}$ can be single-valued even when QVI (1.1) has multiple solutions. On the other hand, single-valuedness of $\mathbf{M_{qvi}}$, locally or globally, requires local or global uniqueness of the solution of QVI (1.1).

(iii) The calmness property can be analyzed using variational analysis tools, such as coderivative criteria for metric subregularity [12, 13]. Estimating the calmness moduli $L_p$ and $L_q$ is possible via coderivatives, but is not straightforward in general.

(iv) Assumptions 4.3 and 4.4 are independent (neither implies the other), as shown by Examples 4.7 and 4.8 below.

Verifying when the assumptions above hold is not a simple task, but we can expect them to be satisfied for nontrivial classes of QVIs, as illustrated by the following example.

**Example 4.6.** Consider affine QVI (1.1), given by

$$F(x) = Ax + b, \quad K(x) = \{z \in \mathbb{R}^n : Cx + Dz = d\},$$

where $A \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n, C, D \in \mathbb{R}^{m\times n}, d \in \mathbb{R}^m$.

Let $A$ and $DA^{-1}D^T$ be nonsingular. This holds, for example, when $A$ is positive definite and $D$ has full rank. If $\bar{x}$ is a solution of the QVI, it satisfies the following KKT system:

$$\begin{cases} A\bar{x} + b + D^T\bar{\lambda} = 0, \\ (C + D)\bar{x} = d, \end{cases} \tag{4.4}$$

where $\bar{\lambda} \in \mathbb{R}^m$ is a Lagrange multiplier associated to the constraints.

Let us examine the parameterized VI (4.3) given by

$$F(x) = Ax + b + p, \quad K(q) = \{z \in \mathbb{R}^n : Cq + Dz = d\}$$

and Assumption 4.3. Solution $x$ of this VI and an associated Lagrange multiplier $\lambda$ satisfy the system

$$\begin{cases} Ax + b + p + D^T\lambda = 0, \\ Cq + Dx = d. \end{cases} \iff \begin{bmatrix} A & D^T \\ D & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -b - p \\ d - Cq \end{bmatrix}. \tag{4.5}$$

In this system, the Schur complement of block $A$ is $-DA^{-1}D^T$. The invertibility assumption ensures the solution is unique, meaning that $\mathbf{M_{vi}}$ is single-valued in the entire domain $\mathbb{R}^n \times \mathbb{R}^n$. Using systems (4.4) and (4.5) together, we have

$$A(x - \bar{x}) + D^T(\lambda - \bar{\lambda}) = -p, \tag{4.6}$$

$$D(x - \bar{x}) = -C(q - \bar{x}). \tag{4.7}$$

Then,

$$x - \bar{x} = A^{-1}(-p - D^T(\lambda - \bar{\lambda})) = -A^{-1}p - A^{-1}D^T(\lambda - \bar{\lambda}). \tag{4.8}$$

Using this in (4.7), we obtain that

$$DA^{-1}p + DA^{-1}D^T(\lambda - \bar{\lambda}) = C(q - \bar{x}),$$

and hence,

$$\lambda - \bar{\lambda} = (DA^{-1}D^T)^{-1}C(q - \bar{x}) - (DA^{-1}D^T)^{-1}DA^{-1}p.$$

Substituting this back into (4.8), we conclude that

$$\begin{aligned} x - \bar{x} &= -A^{-1}p - A^{-1}D^T(DA^{-1}D^T)^{-1}C(q - \bar{x}) + A^{-1}D^T(DA^{-1}D^T)^{-1}DA^{-1}p \\ &= (-A^{-1} + A^{-1}D^T(DA^{-1}D^T)^{-1}DA^{-1})p - A^{-1}D^T(DA^{-1}D^T)^{-1}C(q - \bar{x}). \end{aligned}$$

This implies that

$$\|x - \bar{x}\| \leq L_p\|p\| + L_q\|q - \bar{x}\|,$$

where

$$L_p = \| - A^{-1} + A^{-1}D^T(DA^{-1}D^T)^{-1}DA^{-1}\|, \quad L_q = \|A^{-1}D^T(DA^{-1}D^T)^{-1}C\|.$$

Therefore, Assumption 4.3 is satisfied globally ($V_p = V_q = \mathbb{R}^n$), if

$$\|A^{-1}D^T(DA^{-1}D^T)^{-1}C\| < \frac{1}{2}. \quad \square$$

The next examples illustrate the independence between Assumption 4.3 and Assumption 4.4.

**Example 4.7.** In the QVI (1.2), let $F : \mathbb{R} \to \mathbb{R}$ be defined by $F(x) = x$, and let $K : \mathbb{R} \rightrightarrows \mathbb{R}$ be defined by $K(x) = [x^2, +\infty)$.

Clearly, $0 \in K(0), 1 \in K(1)$, and $N_{K(0)}(0) = N_{K(1)}(1) = (-\infty, 0]$, so that

$$0 \in 0 + (-\infty, 0] = F(0) + N_{K(0)}(0), \text{ and } 0 \in 1 + (-\infty, 0] = F(1) + N_{K(1)}(1).$$

Thus $\bar{x}_1 = 0$ and $\bar{x}_2 = 1$ are solutions of the QVI. Since the QVI solution is not unique, Assumption 4.4 cannot hold.

We now examine Assumption 4.3 at the point $\bar{x}_1 = 0$. We have

$$x \in \mathbf{M_{vi}}(p, q) \iff 0 \in p + x + N_{[q^2, +\infty)}(x).$$

This parametric VI has a unique solution because the mapping $x \mapsto p + x$ is strongly monotone.

If $-p > q^2$, then $x = -p$ is the solution, because

$$0 \in p - p + N_{[q^2, \infty)}(-p) \iff 0 \in 0 + \{0\}.$$

If $-p < q^2$, then $x = q^2$ is the solution, because

$$0 \in p + q^2 + N_{[q^2,+\infty)}(q^2) \iff 0 \in p + q^2 + (-\infty, 0].$$

Therefore,

$$x = \max\{-p, q^2\}, \text{ and hence } |x| \le |p| + |q|^2.$$

Note that

$$|q|^2 \le \frac{1}{3}|q| \iff 3|q|^2 - |q| \le 0 \iff |q| \in \Big[0, \frac{1}{3}\Big] \iff q \in \Big[-\frac{1}{3}, \frac{1}{3}\Big].$$

Thus, for $(p, q) \in \mathbb{R} \times [-\frac{1}{3}, \frac{1}{3}]$,

$$|\mathbf{M_{vi}}(p, q) - \mathbf{M_{vi}}(0, 0)| = |x| \le |p| + |q|^2 \le |p| + \frac{1}{3}|q|,$$

and Assumption 4.3 is satisfied with $L_p = 1$, and $L_q = \frac{1}{3} < \frac{1}{2}$.

**Example 4.8.** In the QVI (1.2), let $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be defined by $K(x) = \{-x\}$, and let $F : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $F(x) = 0$.

Any solution $\bar{x}$ of the QVI satisfies

$$0 \in 0 + N_{\{-\bar{x}\}}(\bar{x}) \implies \bar{x} = -\bar{x} \implies \bar{x} = 0.$$

For the parametric QVI,

$$\begin{aligned} 0 \in \mathbf{M_{qvi}}(p, q) \implies 0 \in p + N_{K(x+q)}(x) &\implies x = -(x + q) \implies x = -\frac{q}{2}, \\ &\implies L_p = 0, L_q = \frac{1}{2} < 1. \end{aligned}$$

Thus, Assumption 4.4 holds. We next show that Assumption 4.3 does not.

For the parametric VI,

$$\begin{aligned} 0 \in \mathbf{M_{vi}}(p, q) \implies 0 \in p + N_{K(q)}(x) &\implies x = -q, \\ &\implies L_p = 0, L_q = 1 > \frac{1}{2}. \end{aligned}$$

Recall that $L_q < \frac{1}{2}$ is required in Assumption 4.3.

## 4.2 Calmness ensures semimonotonicity

We now show that $\mathbf{M_{vi}}$ or $\mathbf{M_{qvi}}$-calmness ensure the lifted mapping $\mathbf{T}$ is semimonotone.

**Theorem 4.9** (Semimonotonicity of QVI mapping)**.** *Under Assumption 4.3 or Assumption 4.4, there exist $\mu^\perp, \mu > 0$, with $\mu^\perp \mu < 1$, and $\mathbf{U} \subset \mathbb{R}^{2n} \times \mathbb{R}^{2n}$ such that the mapping $\mathbf{T}$ in* (2.2) *is $(\mu^\perp, \mu)$-semimonotone at $(\bar{\mathbf{x}}, \mathbf{0})$ on $\mathbf{U}$.*

*Proof.* First, suppose Assumption 4.3 holds. Take $x_1, x_2 \in V_q$ and $y_1 \in V_p$ with $y_1 \in F(x_1) + N_{K(x_2)}(x_1)$, meaning that $x_1 = \mathbf{M_{vi}}(-y_1, x_2)$. Then

$$\|x_1 - \bar{x}\| = \|\mathbf{M_{vi}}(-y_1, x_2) - \mathbf{M_{vi}}(0, \bar{x})\| \le L_p\|y_1\| + L_q\|x_2 - \bar{x}\|. \tag{4.9}$$

Therefore,

$$\|x_2 - \bar{x}\| \le \|x_2 - x_1\| + \|x_1 - \bar{x}\| \le \|x_2 - x_1\| + L_p\|y_1\| + L_q\|x_2 - \bar{x}\|,$$

implying that

$$\|x_2 - \bar{x}\| \le \frac{L_p}{1 - L_q}\|y_1\| + \frac{1}{1 - L_q}\|x_2 - x_1\|. \tag{4.10}$$

Using the Cauchy-Schwarz inequality and (4.9), we have that

$$\langle y_1, x_1 - \bar{x}\rangle \ge -\|y_1\|\|x_1 - \bar{x}\| \ge -\|y_1\|(L_p\|y_1\| + L_q\|x_2 - \bar{x}\|) = -L_p\|y_1\|^2 - L_q\|y_1\|\|x_2 - \bar{x}\|.$$

Using (4.10), we conclude that

$$\begin{aligned}
\langle y_1, x_1 - \bar{x}\rangle &\geq -L_p\|y_1\|^2 - L_q\|y_1\|\left(\frac{L_p}{1-L_q}\|y_1\| + \frac{1}{1-L_q}\|x_2 - x_1\|\right), \\
&= -\left(L_p + \frac{L_pL_q}{1-L_q}\right)\|y_1\|^2 - \frac{L_q}{1-L_q}\|y_1\|\|x_2 - x_1\|, \\
&= -\frac{L_p}{1-L_q}\|y_1\|^2 - \frac{L_q}{1-L_q}\|y_1\|\|x_1 - x_2\|. \qquad (4.11)
\end{aligned}$$

Having in mind the property (4.2), denoting

$$Q(x_1, x_2, y_1) = \langle y_1, x_1 - \bar{x}\rangle + \frac{\mu^\perp}{2}\|x_1 - x_2\|^2 + \frac{\mu}{2}\|y_1\|^2,$$

we need to show $Q(x_1, x_2, y_1) \geq 0$. Using (4.11), we obtain that

$$Q(x_1, x_2, y_1) \geq \left(\frac{\mu}{2} - \frac{L_p}{1-L_q}\right)\|y_1\|^2 - \frac{L_q}{1-L_q}\|y_1\|\|x_1 - x_2\| + \frac{\mu^\perp}{2}\|x_1 - x_2\|^2.$$

Therefore, $Q(x_1, x_2, y_1) \geq 0$ would hold if the following matrix is positive semidefinite:

$$\mathcal{H}_1 = \begin{bmatrix} \frac{\mu}{2} - \frac{L_p}{1-L_q} & \frac{-L_q}{2(1-L_q)} \\ \frac{-L_q}{2(1-L_q)} & \frac{\mu^\perp}{2} \end{bmatrix} \succeq 0.$$

Note that

$$\mathcal{H}_1 \succeq 0 \iff \begin{cases} \mu > \frac{2L_p}{1-L_q}, \\ \mu^\perp > 0, \\ \det \mathcal{H}_1 = \frac{\mu\mu^\perp}{4} - \frac{\mu^\perp L_p}{2(1-L_q)} - \frac{L_q^2}{4(1-L_q)^2} \geq 0. \end{cases}$$

Since $0 < L_q < 1/2$, we have $\frac{L_q^2}{(1-L_q)^2} < 1$. Taking $0 < \varepsilon < 1 - \frac{L_q^2}{(1-L_q)^2}$, we can choose

$$0 < \mu^\perp < \frac{1-L_q}{2L_p}\left(1 - \varepsilon - \frac{L_q^2}{(1-L_q)^2}\right) \text{ and } \mu = \frac{1-\varepsilon}{\mu^\perp} > 0. \qquad (4.12)$$

For these parameters,

$$\mu = (1-\varepsilon)\frac{1}{\mu^\perp} > (1-\varepsilon)\frac{2L_p}{(1-L_q)(1-\varepsilon - \frac{L_q^2}{(1-L_q)^2})} > \frac{2L_p}{1-L_q}.$$

We have that $\mu^\perp\mu = 1 - \varepsilon < 1$. Finally,

$$\begin{aligned}
\det \mathcal{H}_1 &= \frac{1-\varepsilon}{4} - \mu^\perp \frac{L_p}{2(1-L_q)} - \frac{L_q^2}{4(1-L_q)^2} \\
&> \frac{1-\varepsilon}{4} - \frac{1-L_q}{2L_p}\left(1 - \varepsilon - \frac{L_q^2}{(1-L_q)^2}\right)\frac{L_p}{2(1-L_q)} - \frac{L_q^2}{4(1-L_q)^2} = 0.
\end{aligned}$$

Since $Q(x_1, x_2, y_1) \geq 0$ when $x_1, x_2 \in V_q$ and $y_1 \in V_p$, we conclude that $\mathbf{T}$ is $(\mu^\perp, \mu)$-semimonotone at $(\bar{\mathbf{x}}, \mathbf{0})$ on $\mathbf{U} = (V_q \times V_q) \times (V_p \times \mathbb{R}^n)$.

Now suppose Assumption 4.4 holds. We can choose $\delta > 0$ such that $\mathbb{B}_\delta \subset V_p$. Taking $y_1 \in V_p$ and $x_1, x_2 \in \{\bar{x}\} + \frac{1}{2}\mathbb{B}_\delta$, we have that $\|x_1 - x_2\| \leq \|x_1 - \bar{x}\| + \|\bar{x} - x_2\| \leq \delta$, implying that $x_1 - x_2 \in V_p$. Therefore,

$$\|x_1 - \bar{x}\| = \|\mathbf{M_{qvi}}(-y_1, x_2 - x_1) - \mathbf{M_{qvi}}(0, 0)\| \leq L_p\|y_1\| + L_q\|x_2 - x_1\|.$$

We obtain that

$$\langle y_1, x_1 - \bar{x}\rangle \geq -L_p\|y_1\|^2 - L_q\|y_1\|\|x_1 - x_2\|,$$

$$Q(x_1, x_2, y_1) \geq \left(\frac{\mu}{2} - L_p\right)\|y_1\|^2 - L_q\|y_1\|\|x_1 - x_2\| + \frac{\mu^\perp}{2}\|x_1 - x_2\|^2.$$

Therefore, $Q(x_1, x_2, y_1) \geq 0$ if

$$\mathcal{H}_2 = \begin{bmatrix} \frac{\mu}{2} - L_p & -\frac{L_q}{2} \\ -\frac{L_q}{2} & \frac{\mu^\perp}{2} \end{bmatrix} \succeq 0.$$

Note that

$$\mathcal{H}_2 \succeq 0 \iff \begin{cases} \mu > 2L_p, \\ \mu^\perp > 0, \\ \det \mathcal{H}_2 = \frac{\mu\mu^\perp}{4} - \frac{\mu^\perp L_p}{2} - \frac{L_q^2}{4} \geq 0. \end{cases}$$

Since $0 \leq L_q < 1$ we can take $\varepsilon < 1 - L_q^2$, $0 < \mu^\perp < \frac{1-L_q^2-\varepsilon}{2L_p}$, $\mu = \frac{1-\varepsilon}{\mu^\perp}$, so that

$$\mu^\perp \mu = 1 - \varepsilon < 1,$$

$$\mu = (1-\varepsilon)\frac{1}{\mu^\perp} > (1-\varepsilon)\frac{2L_p}{1 - L_q^2 - \varepsilon} \geq 2L_p.$$

$$\det \mathcal{H}_2 = \frac{1-\varepsilon}{4} - \frac{\mu^\perp L_p}{2} - \frac{L_q^2}{4} > \frac{1-\varepsilon}{4} - \frac{1 - L_q^2 - \varepsilon}{2L_p}\frac{L_p}{2} - \frac{L_q^2}{4} = 0.$$

We conclude that $Q(x_1, x_2, y_1) \geq 0$ when $x_1, x_2 \in \{\bar{x}\} + \frac{1}{2}\mathbb{B}_\delta$ and $y_1 \in V_p$. Thus, $\mathbf{T}$ is $(\mu^\perp, \mu)$-semimonotone at $(\bar{\mathbf{x}}, \mathbf{0})$ on $\mathbf{U} = ((\{\bar{x}\} + \frac{1}{2}\mathbb{B}_\delta) \times (\{\bar{x}\} + \frac{1}{2}\mathbb{B}_\delta)) \times (V_p \times \mathbb{R}^n)$. □

# 5 Global Convergence Analysis

When Assumption 4.3 or Assumption 4.4 hold on the whole domain, we show in Theorem 5.2 below that Algorithm 2 converges globally. Independently, under appropriate conditions, global convergence with linear rate is proven for the case of QVIs with moving set, in Theorem 5.6.

## 5.1 Global convergence under global calmness

Before stating our result, we revisit Example 4.1, to show that the mapping $\mathbf{T}$ therein satisfies the semimonotone inequality in Definition 4.2 in the whole space (even though neither monotonicity, even local, nor elicitability hold for $\mathbf{T}$).

**Example 5.1** (Continuation of Example 4.1)**.** For the same mappings $F$ and $K$ in Example 4.1, we shall show that $\mathbf{T}$ is semimonotone at $(\bar{\mathbf{x}}, \mathbf{0}) = (\mathbf{0}, \mathbf{0})$ in two different ways.

First, we show semimonotonicity directly using the concept in (4.2). We have that

$$y_1 \in F(x_1) + N_{K(x_2)}(x_1) \implies x_1 = tx_2, \text{ for some } t \in \mathbb{R}, \text{ and } y_1 = x_1 + v, \text{ with } v \in [x_2]^\perp.$$

Then $y_1 = tx_2 + v$, with $v \in [x_2]^\perp$, and

$$\begin{aligned}\langle y_1, x_1 - 0\rangle + \frac{\mu^\perp}{2}\|x_1 - x_2\|^2 + \frac{\mu}{2}\|y_1\|^2 &= \langle tx_2 + v, tx_2\rangle + \frac{\mu^\perp}{2}\|tx_2 - x_2\|^2 + \frac{\mu}{2}\|tx_2 + v\|^2 \\ &= t^2\|x_2\|^2 + \frac{\mu^\perp(t-1)^2}{2}\|x_2\|^2 + \frac{\mu t^2}{2}\|x_2\|^2 + \frac{\mu}{2}\|v\|^2 \geq 0,\end{aligned}$$

for any choice of $\mu \geq 0, \mu^\perp \geq 0$.

Alternatively, we can verify that Assumption 4.3 holds, and apply Theorem 4.9. Let $x \in \mathbf{M_{vi}}(p, q)$ be a solution for the parametrized VI in question. Then

$$0 \in p + F(x) + N_{K(q)}(x) \iff x = tq, \text{for some } t \in \mathbb{R}, \text{and } 0 = p + x + v, \text{with } v \in [q]^\perp.$$

If $q = 0$, we would have $x = 0$, and then $\|\mathbf{M_{vi}}(p, 0) - \mathbf{M_{vi}}(0, 0)\| = 0 \leq L_q\|q\| + L_p\|p\|$, for any $L_p, L_q \geq 0$. So assume $q \neq 0$. Since $-p = tq + v$, taking the inner product with $q$ we obtain that

$$-\langle p, q\rangle = t\|q\|^2 \implies t = -\frac{\langle p, q\rangle}{\|q\|^2} \implies x = -\frac{\langle p, q\rangle}{\|q\|^2}q.$$

Therefore, the solution $x$ is unique and

$$\|\mathbf{M_{vi}}(p,q) - \mathbf{M_{vi}}(0,0)\| = \left\| -\frac{\langle p,q\rangle}{\|q\|^2} q \right\| = \frac{|\langle p,q\rangle|}{\|q\|} \le \frac{\|p\|\,\|q\|}{\|q\|} = \|p\|\,.$$

Writing the right hand side as $\|p\| = L_q\|q\| + L_p\|p\|$, Assumption 4.3 holds with $L_q = 0 < 1/2$, $L_p = 1$. □

We now prove that our QVI progressive decoupling+ method is globally convergent.

**Theorem 5.2** (Global convergence of Algorithm 2)**.** *Assume that* $\operatorname{dom} K = \mathbb{R}^n$*,* $F$ *is continuous and either* $K$ *is compact-valued or* $F$ *is monotone. Consider parameters in Algorithm 2 satisfying*

$$r \in (\mu^\perp, 1/\mu), \lambda_x \in (0, 2(1 - r\mu)), \lambda_y \in (0, 2(1 - \mu^\perp/r)), \tag{5.1}$$

*where* $\mu, \mu^\perp > 0$ *with* $\mu\mu^\perp < 1$ *are given by Theorem 4.9. Then the following statements hold:*

*(i) If Assumption 4.3 holds with* $V_p = V_q = \mathbb{R}^n$*, then either a solution of QVI* (1.1) *is reached in a finite number of iterations, or the generated sequence* $\{x^k\}$ *is bounded and every accumulation point of this sequence is a solution of QVI* (1.1)*. In particular, if* $\bar{x}$ *is the unique solution of QVI* (1.1)*, then the sequence converges to* $\bar{x}$*.*

*(ii) If Assumption 4.4 holds with* $V_p = V_q = \mathbb{R}^n$*, then the generated sequence* $\{x^k\}$ *reaches a solution in a finite number of iterations or converges to the (unique in this case) QVI solution* $\bar{x}$*.*

*Proof.* Global convergence of progressive decoupling+ depends on three properties stated in Assumption IV of [15], referred to as A1, A2, and A3. We next show that these hold in our case.

Property A1 is semimonotonicity, which follows from Theorem 4.9 with $\mathbf{U} = \mathbb{R}^{2n} \times \mathbb{R}^{2n}$, under Assumption 4.3 or under Assumption 4.4.

Property A2 concerns outer semicontinuity of the mapping $\mathbf{T}$, which holds here on the entire domain. To see this, let

$$\left( \begin{bmatrix} x_1^i \\ x_2^i \end{bmatrix}, \begin{bmatrix} y_1^i \\ y_2^i \end{bmatrix} \right) \in \operatorname{gph} \mathbf{T} \text{ be such that } \left( \begin{bmatrix} x_1^i \\ x_2^i \end{bmatrix}, \begin{bmatrix} y_1^i \\ y_2^i \end{bmatrix} \right) \to \left( \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix} \right), \quad \text{as } i \to \infty.$$

Clearly, we have that $y_1^i \in F(x_1^i) + N_{K(x_2^i)}(x_1^i)$ and $y_2^i = 0$. Thus $\bar{y}_2 = \lim_{i\to\infty} y_2^i = 0$. For any $z \in K(\bar{x}_2)$, by the inner semicontinuity of $K(\cdot)$, we can choose $z^i \in K(x_2^i)$ with $z^i \to z$. Since $y_1^i - F(x_1^i) \in N_{K(x_2^i)}(x_1^i)$, it holds that $\langle y_1^i - F(x_1^i), z^i - x_1^i\rangle \le 0$. By the continuity of $F(\cdot)$, taking the limit as $i \to \infty$, we obtain that

$$\langle \bar{y}_1 - F(\bar{x}_1), z - \bar{x}_1\rangle \le 0.$$

Since $z \in K(\bar{x}_2)$ was arbitrary, it follows that $\bar{y}_1 \in F(\bar{x}_1) + N_{K(\bar{x}_2)}(\bar{x}_1)$. Therefore,

$$\left( \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \end{bmatrix} \right) \in \operatorname{gph} \mathbf{T}.$$

Hence, $\operatorname{gph} \mathbf{T}$ is closed, and therefore $\mathbf{T}$ is outer semicontinuous.

Finally, property A3 requires the resolvent $\mathcal{J}_{\frac{1}{r}\mathbf{T}}$ to have full domain. Indeed, for $\mathbf{x} \in \mathbb{R}^{2n}$,

$$\begin{aligned} \mathbf{y} \in \mathcal{J}_{\frac{1}{r}\mathbf{T}}(\mathbf{x}) \quad &\iff \quad \mathbf{x} \in \mathbf{y} + \frac{1}{r}\mathbf{T}(\mathbf{y}) \iff r(\mathbf{x} - \mathbf{y}) \in \mathbf{T}(\mathbf{y}) \\ &\iff \quad \begin{cases} r(x_1 - y_1) \in F(y_1) + N_{K(y_2)}(y_1), \\ r(x_2 - y_2) = 0. \end{cases} \end{aligned}$$

Or equivalently, $y_2 = x_2$ and $y_1$ solves $\mathrm{VI}(F + rI - rx_1, K(x_2))$. Since $x_2 \in \operatorname{dom} K$, we have that $K(x_2) \neq \emptyset$. If $K$ is compact-valued, then $K(x_2)$ is compact, and existence of a solution to $\mathrm{VI}(F + rI - rx_1, K(x_2))$ follows from the continuity of $F$. If $F$ is monotone, then $F + rI - rx_1$ is strongly monotone, and therefore the VI has the unique solution. In either case, $\mathcal{J}_{\frac{1}{r}\mathbf{T}}(\mathbf{x}) \neq \emptyset$ for all $\mathbf{x} \in \mathbb{R}^{2n}$. Thus, $\mathcal{J}_{\frac{1}{r}\mathbf{T}}$ has full domain.

When the parameters lie in the nonempty sets given by (5.1), Corollary 4.10 of [15] ensures that Algorithm 2 either reaches a solution in a finite number of iterations or generates a bounded sequence whose accumulation points are solutions of QVI 1.1 (see also item (iii) of Theorem 4.9 in [15]). If the solution is unique, then the sequence converges to this solution. The case of Assumption 4.4 falls into the uniqueness setting. □

Convergence of Algorithm 1 follows taking $\lambda_x = \lambda_y = 1$ in (5.1).

**Corollary 5.3** (Global convergence of Algorithm 1 with fixed proximal parameter)**.** *Consider the QVI progressive decoupling method given in Algorithm 1. Let for all iterations, $r_k = r \in (\mu^{\perp}, 1/\mu)$. Then, if the mappings $F$ and $K$ are as in Theorem 5.2, under the assumptions in (i) and (ii) therein, the convergence properties of Algorithm 1 with a fixed proximal parameter are the same as stated in Theorem 5.2 for Algorithm 2.* □

## 5.2 The moving set case

For the class of QVIs whose mapping $K$ represents a moving set: $K(x) = c(x) + C$, where $C \subset \mathbb{R}^n$ is a fixed closed convex set and $c : \mathbb{R}^n \to \mathbb{R}^n$ is a given function, we next establish convergence of Algorithm 1 directly, allowing variable proximal parameters, without invoking the framework of [15]. Moreover, under suitable conditions our algorithm converges with the global linear rate.

Formulations of QVI with a moving set appear in numerous mechanical applications with implicit-obstacle-type constraints, as described in [28]. This specific structure has been extensively studied in the QVI literature; see [18]. A proximal point method for QVIs with moving set is proposed in [30].

We state the following proposition, which will be used in the sequel. It is an immediate consequence of some results in [33]. Specifically, [33, Lemma 2] on a certain property of projections onto a moving set, and [33, Corollary 2] on existence of solutions for a general QVI which requires a condition on projections (satisfied under appropriate assumptions in the moving set case).

**Proposition 5.4.** *For the QVI* (1.1)*, let $F$ be strongly monotone with modulus $\sigma > 0$ and Lipschitz continuous with modulus $\ell_F > 0$. Let $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be given by $K(x) = c(x) + C$, where $c : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous with modulus $\ell_c > 0$, and $C \subset \mathbb{R}^n$ is a closed, convex set. If $\ell_c < \frac{\sigma}{\ell_F}$, then the QVI is solvable, and the solution is unique.*

We now introduce an assumption required to establish global convergence of Algorithm 1 for the moving set case.

**Assumption 5.5** (Parameters relations for a QVI with moving set)**.** *Let QVI* (1.1) *be such that:*

1. *$F$ is $\sigma$-strongly monotone and Lipschitz continuous with modulus $\ell_F > 0$;*
2. *$K(x) = c(x) + C$, where $C \subset \mathbb{R}^n$ is a closed convex set and $c : \mathbb{R}^n \to \mathbb{R}^n$ is Lipschitz continuous with modulus $\ell_c > 0$;*
3. *For some $\tau \in (\frac{1}{2}, 1)$,*

$$2\sigma - \ell_F^2 > 0 \quad \text{and} \quad \ell_c \leq (\sqrt{\tau} - \frac{\sqrt{2}}{2})\sqrt{2\sigma - \ell_F^2}. \quad \square$$

Before stating the next theorem, note that $\sqrt{2\tau} - \tau \in (0, \frac{1}{2})$ because

$$\begin{cases} \sqrt{2\tau} - \tau = \sqrt{2}\sqrt{\tau} - \tau \geq \sqrt{2}\tau - \tau > 0, \\ \sqrt{2\tau} - \tau = \sqrt{2}\sqrt{\tau} - \tau \leq \max\limits_{u \in (\frac{\sqrt{2}}{2}, 1)} \{\sqrt{2}u - u^2\} < \sqrt{2}\left(\frac{\sqrt{2}}{2}\right) - \left(\frac{\sqrt{2}}{2}\right)^2 = \frac{1}{2}. \end{cases}$$

Now, we can state the following result.

**Theorem 5.6.** *[Convergence of Algorithm 1 with varying proximal parameter for moving set QVIs] Under Assumption 5.5, the QVI* (1.1) *has the unique solution $\bar{x}$, and from any starting point, Algorithm 1 generates a sequence $\{x^k\}$ converging to this solution, provided the parameters $r_k$ satisfy $0 < r_0 \leq r_k \leq r_{k+1} \leq \left(\frac{\frac{1}{2} - (\sqrt{2\tau} - \tau)}{\frac{1}{2} + (\sqrt{2\tau} - \tau)}\right)(2\sigma - \ell_F^2)$, for all $k$.*

*Moreover, the primal-dual sequence* $\{\|x^k - \bar{x}\|^2 + \frac{1}{r_k^2}\|y^k\|^2\}$ *converges to zero with a global linear rate.*

*Proof.* First, we establish the existence and uniqueness of the QVI solution. Note that $\frac{\sigma}{\ell_F} > 0$, and

$$\begin{aligned}\frac{\sqrt{2\sigma - \ell_F^2}}{\frac{\sigma}{\ell_F}} &= \sqrt{2\ell_F^2 \frac{1}{\sigma} - \ell_F^4 \frac{1}{\sigma^2}} \\ &\leq \max_{\frac{1}{\sigma}>0} \sqrt{2\ell_F^2\left(\frac{1}{\sigma}\right) - \ell_F^4\left(\frac{1}{\sigma}\right)^2} \\ &= \sqrt{2\ell_F^2\left(\frac{1}{\ell_F^2}\right) - \ell_F^4\left(\frac{1}{\ell_F^2}\right)^2} = 1.\end{aligned}$$

By the inequality above and Assumption 5.5, it follows that $\ell_c \leq (\sqrt{\tau} - \frac{\sqrt{2}}{2})\frac{\sigma}{\ell_F} < \frac{\sigma}{\ell_F}$. Proposition 5.4 ensures the existence and uniqueness of the QVI solution.

Let $\bar{x}$ be the solution of the QVI. By the definition of the iterates of Algorithm 1 and since $0 < r_k \leq r_{k+1}$, we obtain that

$$\begin{aligned}\|x^{k+1} - \bar{x}\|^2 + \frac{1}{r_{k+1}^2}\|y^{k+1}\|^2 &\leq \|x^{k+1} - \bar{x}\|^2 + \frac{1}{r_k^2}\| - \frac{r_k}{2}(\hat{x}^k - x^k - \frac{1}{r_k}y^k)\|^2 \\ &= \|x^{k+1} - \bar{x}\|^2 + \|\frac{1}{2}(\hat{x}^k + x^k - \frac{1}{r_k}y^k) - x^k\|^2 \\ &= \|x^{k+1} - \bar{x}\|^2 + \|x^{k+1} - x^k\|^2 \\ &= \frac{1}{2}\left(\|(x^{k+1} - \bar{x}) + (x^{k+1} - x^k)\|^2 + \|(x^{k+1} - \bar{x}) - (x^{k+1} - x^k)\|^2\right) \\ &= \frac{1}{2}\left(\|2x^{k+1} - x^k - \bar{x}\|^2 + \|x^k - \bar{x}\|^2\right) \\ &= \frac{1}{2}\left(\|\hat{x}^k - \frac{1}{r_k}y^k - \bar{x}\|^2 + \|x^k - \bar{x}\|^2\right) \\ &= \frac{1}{2}(\|\hat{x}^k - \bar{x}\|^2 + \|x^k - \bar{x}\|^2 + \frac{1}{r_k^2}\|y^k\|^2 - 2\langle \hat{x}^k - \bar{x}, \frac{1}{r_k}y^k\rangle) \\ &= \frac{1}{2}(\|x^k - \bar{x}\|^2 + \frac{1}{r_k^2}\|y^k\|^2) + \frac{1}{2}\|\hat{x}^k - \bar{x}\|^2 - \langle \hat{x}^k - \bar{x}, \frac{1}{r_k}y^k\rangle.\end{aligned}$$

We next analyze the last term in the right-hand side of the relation above. Note that for any $\theta_k \neq 0$ (this $\theta_k$ is introduced here for a reason that will become clear in the sequel), it holds that

$$-\langle \hat{x}^k - \bar{x}, \frac{1}{r_k}y^k\rangle = -\langle \theta_k(\hat{x}^k - \bar{x}), \frac{1}{\theta_k r_k}y^k\rangle \leq \frac{\theta_k^2}{2}\|\hat{x}^k - \bar{x}\|^2 + \frac{1}{2\theta_k^2 r_k^2}\|y^k\|^2.$$

Combining the relations above, we obtain that

$$\|x^{k+1} - \bar{x}\|^2 + \frac{1}{r_{k+1}^2}\|y^{k+1}\|^2 \leq \frac{1}{2}\|x^k - \bar{x}\|^2 + \frac{1+\theta_k^2}{2}\|\hat{x}^k - \bar{x}\|^2 + \left(\frac{1}{2} + \frac{1}{2\theta_k^2}\right)\frac{1}{r_k^2}\|y^k\|^2. \quad (5.2)$$

The task now is to estimate the middle term in the right-hand side of (5.2), so that the expression reduces to a suitable multiple of $\|x^k - \bar{x}\|^2 + \frac{1}{r_k^2}\|y^k\|^2$, implying (linear) convergence of the corresponding sequence.

Observe that

$$\bar{x} \in K(\bar{x}) = c(\bar{x}) + C \implies \bar{x} - c(\bar{x}) + c(x^k - \frac{1}{r_k}y^k) \in K(x^k - \frac{1}{r_k}y^k) = K^k.$$

Therefore, the fact that $\hat{x}^k$ solves $\mathrm{VI}(F + r_k I - r_k x^k - y^k, K^k)$ implies that

$$\langle F(\hat{x}^k) + r_k\hat{x}^k - r_k x^k - y^k, \bar{x} - c(\bar{x}) + c(x^k - \frac{1}{r_k}y^k) - \hat{x}^k\rangle \geq 0. \quad (5.3)$$

On the other hand,

$$\hat{x}^k \in K(x^k - \frac{1}{r_k}y^k) = c(x^k - \frac{1}{r_k}y^k) + C \implies \hat{x}^k - c(x^k - \frac{1}{r_k}y^k) + c(\bar{x}) \in K(\bar{x}).$$

The fact that $\bar{x}$ solves QVI$(F, K(\cdot))$ implies that

$$\langle F(\bar{x}), \hat{x}^k - c(x^k - \frac{1}{r_k}y^k) + c(\bar{x}) - \bar{x}\rangle \geq 0,$$

or equivalently,

$$\langle -F(\bar{x}), \bar{x} - c(\bar{x}) + c(x^k - \frac{1}{r_k}y^k) - \hat{x}^k\rangle \geq 0.$$

Adding the last inequality above with (5.3), we obtain that

$$\langle F(\hat{x}^k) - F(\bar{x}) + r_k\hat{x}^k - r_k x^k - y^k, \bar{x} - \hat{x}^k + c(x^k - \frac{1}{r_k}y^k) - c(\bar{x})\rangle \geq 0.$$

Hence, by the strong monotonicity of $F$, we obtain that

$$\begin{aligned}
\sigma\|\hat{x}^k - \bar{x}\|^2 &\leq \langle F(\hat{x}^k) - F(\bar{x}), \hat{x}^k - \bar{x}\rangle \\
&\leq \langle F(\hat{x}^k) - F(\bar{x}), c(x^k - \frac{1}{r_k}y^k) - c(\bar{x})\rangle + r_k\langle \hat{x}^k - x^k - \frac{1}{r_k}y^k, \bar{x} - \hat{x}^k\rangle \\
&\quad + r_k\langle \hat{x}^k - x^k - \frac{1}{r_k}y^k, c(x^k - \frac{1}{r_k}y^k) - c(\bar{x})\rangle. \qquad (5.4)
\end{aligned}$$

We next estimate the three terms in the right-hand side of (5.4).

First, by the Lipschitz continuity of $F$ and of $c$, we have that

$$\begin{aligned}
\langle F(\hat{x}^k) - F(\bar{x}), c(x^k - \frac{1}{r_k}y^k) - c(\bar{x})\rangle &\leq \frac{1}{2}\|F(\hat{x}^k) - F(\bar{x})\|^2 + \frac{1}{2}\|c(x^k - \frac{1}{r_k}y^k) - c(\bar{x})\|^2 \\
&\leq \frac{\ell_F^2}{2}\|\hat{x}^k - \bar{x}\|^2 + \frac{\ell_c^2}{2}\|x^k - \frac{1}{r_k}y^k - \bar{x}\|^2. \qquad (5.5)
\end{aligned}$$

Second, it holds that

$$r_k\langle \hat{x}^k - x^k - \frac{1}{r_k}y^k, \bar{x} - \hat{x}^k\rangle = \frac{r_k}{2}(\|\bar{x} - x^k - \frac{1}{r_k}y^k\|^2 - \|\hat{x}^k - x^k - \frac{1}{r_k}y^k\|^2 - \|\hat{x}^k - \bar{x}\|^2). \quad (5.6)$$

Third, using again the Lipschitz continuity of $c$, we obtain that

$$r_k\langle \hat{x}^k - x^k - \frac{1}{r_k}y^k, c(x^k - \frac{1}{r_k}y^k) - c(\bar{x})\rangle \leq \frac{r_k}{2}\|\hat{x}^k - x^k - \frac{1}{r_k}y^k\|^2 + \frac{r_k\ell_c^2}{2}\|x^k - \frac{1}{r_k}y^k - \bar{x}\|^2. \quad (5.7)$$

Combining now (5.5)–(5.7) with (5.4), we obtain that

$$\sigma\|\hat{x}^k - \bar{x}\|^2 \leq \frac{\ell_F^2 - r_k}{2}\|\hat{x}^k - \bar{x}\|^2 + \frac{\ell_c^2(r_k + 1)}{2}\|x^k - \bar{x} - \frac{1}{r_k}y^k\|^2 + \frac{r_k}{2}\|x^k - \bar{x} + \frac{1}{r_k}y^k\|^2.$$

As $r_k + 2\sigma - \ell_F^2 > 0$, it follows that

$$\|\hat{x}^k - \bar{x}\|^2 \leq \frac{\ell_c^2(r_k + 1)}{r_k + 2\sigma - \ell_F^2}\|x^k - \bar{x} - \frac{1}{r_k}y^k\|^2 + \frac{r_k}{r_k + 2\sigma - \ell_F^2}\|x^k - \bar{x} + \frac{1}{r_k}y^k\|^2.$$

Denoting $\gamma_k = \max\{r_k, \ell_c^2(r_k + 1)\}$, it then holds that

$$\begin{aligned}
\|\hat{x}^k - \bar{x}\|^2 &\leq \frac{\gamma_k}{r_k + 2\sigma - \ell_F^2}(\|x^k - \bar{x} - \frac{1}{r_k}y^k\|^2 + \|x^k - \bar{x} + \frac{1}{r_k}y^k\|^2) \\
&= \frac{2\gamma_k}{r_k + 2\sigma - \ell_F^2}(\|x^k - \bar{x}\|^2 + \frac{1}{r_k^2}\|y^k\|^2).
\end{aligned}$$

Using this inequality in (5.2), we obtain that

$$\begin{aligned}
\|x^{k+1}-\bar{x}\|^2+\frac{1}{r_{k+1}^2}\|y^{k+1}\|^2 &\leq \left(\frac{1}{2}+\frac{(1+\theta_k^2)\gamma_k}{(r_k+2\sigma-\ell_F^2)}\right)\|x^k-\bar{x}\|^2 \\
&\quad+\left(\frac{1}{2}+\frac{1}{2\theta_k^2}+\frac{(1+\theta_k^2)\gamma_k}{(r_k+2\sigma-\ell_F^2)}\right)\frac{1}{r_k^2}\|y^k\|^2 \\
&\leq \left(\frac{1}{2}+\frac{1}{2\theta_k^2}+\frac{(1+\theta_k^2)\gamma_k}{(r_k+2\sigma-\ell_F^2)}\right)(\|x^k-\bar{x}\|^2+\frac{1}{r_k^2}\|y^k\|^2).
\end{aligned}$$

To establish linear convergence of the sequence $\{\|x^k-\bar{x}\|^2+\frac{1}{r_k^2}\|y^k\|^2\}$ to zero, we have to show that the multiple in the right-hand side of the last relation is bounded above by some $t<1$ for all $k$.

Take $\theta_k=\left(\frac{r_k+2\sigma-\ell_F^2}{2\gamma_k}\right)^{\frac{1}{4}}$ and denote $\beta_k=\left(\frac{\gamma_k}{r_k+2\sigma-\ell_F^2}\right)^{\frac{1}{2}}$. Note that $\theta_k^2=\frac{1}{\sqrt{2}\beta_k}$. We have that

$$\begin{aligned}
\left(\frac{1}{2}+\frac{1}{2\theta_k^2}+\frac{(1+\theta_k^2)\gamma_k}{(r_k+2\sigma-\ell_F^2)}\right) &= \frac{1}{2}+\frac{\sqrt{2}}{2}\beta_k+\beta_k^2+\frac{1}{\sqrt{2}\beta_k}\beta_k^2 \\
&= \frac{1}{2}+\sqrt{2}\beta_k+\beta_k^2=\left(\beta_k+\frac{\sqrt{2}}{2}\right)^2.
\end{aligned}$$

Hence, the claim would follow if we show, for example, that

$$\left(\beta_k+\frac{\sqrt{2}}{2}\right)^2\leq\tau,\ \text{for all } k, \tag{5.8}$$

where $\tau\in(1/2,1)$ is given in the assumptions of the Theorem. Since $\beta_k>0$, the condition (5.8) is equivalent to

$$\beta_k\leq\sqrt{\tau}-\frac{\sqrt{2}}{2}. \tag{5.9}$$

To conclude the proof, it remains to verify (5.9).

As $0<\sigma\leq\ell_F$ and $2\sigma-\ell_F^2\geq 0$, it holds that

$$2\sigma-\ell_F^2\leq 2\ell_F-\ell_F^2\leq\max_{u\in\mathbb{R}}\{2u-u^2\}=1.$$

Since $2\sigma-\ell_F^2\leq 1$, we have that $r_k(2\sigma-\ell_F^2)+2\sigma-\ell_F^2\leq r_k+2\sigma-\ell_F^2$. Hence,

$$(r_k+1)(2\sigma-\ell_F^2)\leq r_k+2\sigma-\ell_F^2. \tag{5.10}$$

Squaring the second inequality in Assumption 5.5, we obtain that

$$\ell_c^2\leq(\tau-\sqrt{2\tau}+\frac{1}{2})(2\sigma-\ell_F^2).$$

Multiplying the latter by $(r_k+1)$, we have

$$\begin{aligned}
\ell_c^2(r_k+1) &\leq (\tau-\sqrt{2\tau}+\frac{1}{2})(2\sigma-\ell_F^2)(r_k+1), \\
&\leq (\tau-\sqrt{2\tau}+\frac{1}{2})(r_k+2\sigma-\ell_F^2),
\end{aligned} \tag{5.11}$$

where the second inequality is by (5.10). Using the condition on the choice of $r_k$, which is

$$r_k\leq\left(\frac{\frac{1}{2}-(\sqrt{2\tau}-\tau)}{\frac{1}{2}+(\sqrt{2\tau}-\tau)}\right)(2\sigma-\ell_F^2),$$

we have that

$$r_k\left(\frac{1}{2}+(\sqrt{2\tau}-\tau)\right)\leq\left(\frac{1}{2}-(\sqrt{2\tau}-\tau)\right)(2\sigma-\ell_F^2),$$

from where it follows that

$$r_k\left(\frac{1}{2}+(\sqrt{2\tau}-\tau)\right)+r_k\left(\frac{1}{2}-(\sqrt{2\tau}-\tau)\right)\le\left(\frac{1}{2}-(\sqrt{2\tau}-\tau)\right)(2\sigma-\ell_F^2+r_k),$$

and hence,

$$r_k\le(\tau-\sqrt{2\tau}+\frac{1}{2})(r_k+2\sigma-\ell_F^2). \tag{5.12}$$

Since $\gamma_k=\max\{r_k,\ell_c^2(r_k+1)\}$, inequalities (5.11) and (5.12) yield

$$\gamma_k\le(\tau-\sqrt{2\tau}+\frac{1}{2})(r_k+2\sigma-\ell_F^2)\implies\frac{\gamma_k}{r_k+2\sigma-\ell_F^2}\le\left(\sqrt{\tau}-\frac{\sqrt{2}}{2}\right)^2$$
$$\implies\beta_k=\left(\frac{\gamma_k}{r_k+2\sigma-\ell_F^2}\right)^{\frac{1}{2}}\le\sqrt{\tau}-\frac{\sqrt{2}}{2}.$$

This verifies (5.9), thus completing the proof. □

To conclude with the convergence theory for Algorithm 2, we now examine its behavior when the regularity conditions required in Theorem 5.2 only hold near a solution.

# 6 Local Convergence Analysis

To exhibit local convergence of the progressive decoupling+ algorithm, a localization condition (the region in which convergence assertions hold), must be specified. To this end, following [15], we define the following neighborhood of a solution.

$$\mathcal{W}_\varepsilon=\left\{(\mathbf{x},\mathbf{y})\in\mathbf{S}\times\mathbf{S}^\perp:\frac{r}{\lambda_x}\|\mathbf{x}-\bar{\mathbf{x}}\|^2+\frac{1}{r\lambda_y}\|\mathbf{y}-\mathbf{0}\|^2\le\varepsilon^2\right\}. \tag{6.1}$$

Equivalently,

$$(\mathbf{x},\mathbf{y})\in\mathcal{W}_\varepsilon\iff\mathbf{x}=\begin{bmatrix}x\\x\end{bmatrix},\mathbf{y}=\begin{bmatrix}y\\-y\end{bmatrix},\text{ with }\frac{2r}{\lambda_x}\|x-\bar{x}\|^2+\frac{2}{r\lambda_y}\|y\|^2\le\varepsilon^2.$$

In particular,

$$(\mathbf{x},\mathbf{y})\in\mathcal{W}_\varepsilon\implies\mathbf{x}=\begin{bmatrix}x\\x\end{bmatrix},\mathbf{y}=\begin{bmatrix}y\\-y\end{bmatrix},\|x-\bar{x}\|\le\varepsilon\sqrt{\frac{\lambda_x}{2r}}\text{ and }\|y\|\le\varepsilon\sqrt{\frac{r\lambda_y}{2}}. \tag{6.2}$$

In addition to the Assumption 4.3 or Assumption 4.4, we introduce another regularity condition, now on the parameterized VI with a proximal term. That it is independent from Assumption 4.3 or Assumption 4.4, is illustrated in Examples 6.2 and 6.3, shortly in the sequel. This new condition is natural, as the proximal term tends to stabilize solutions near the reference point $\bar{x}$. Given $r>0$ and $\bar{x}$ a solution of the QVI 1.2, we introduce the solution mapping $\mathbf{M^r}:\mathbb{R}^n\times\mathbb{R}^n\rightrightarrows\mathbb{R}^n$,

$$x\in\mathbf{M^r}(p,q)\iff 0\in p+F(x)+r(x-\bar{x})+N_{K(q)}(x). \tag{6.3}$$

**Assumption 6.1** ($\mathbf{M^r}$-calmness)**.** *The mapping $\mathbf{M^r}$ is locally single-valued and calm at $((0,\bar{x}),\bar{x})$: there exist neighborhoods $V'_p\ni 0, V'_q\ni\bar{x}$ and constants $L'_p,L'_q\ge 0$ such that $\mathbf{M^r}$ is single-valued in $V'_p\times V'_q$ and*

$$\|\mathbf{M^r}(p,q)-\mathbf{M^r}(0,\bar{x})\|\le L'_p\|p\|+L'_q\|q-\bar{x}\|,\ \textit{for all } (p,q)\in V'_p\times V'_q.$$

Note that we do not require bounds for the calmness constants in Assumption 6.1. In this case, single-valuedness follows from strong monotonicity induced by the proximal term, as guaranteed by Theorem 2.3.3-(b) from [19]. Subregularity is more easily satisfied, in particular because there are no bounds required for the calmness constants. Results from [12, 13] can be useful to analyze subregularity further.

Assumption 6.1 is independent of Assumptions 4.3 and 4.4, as illustrated by the following examples.

**Example 6.2.** In the QVI (1.2), let $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be the constant mapping $K(x) = \mathbb{R}^n$ and let $F : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $F(x) = -x$.

If $\bar{x}$ solves the QVI, then

$$0 \in -\bar{x} + \{0\} \implies \bar{x} = 0.$$

Analyzing the parametric QVI,

$$0 \in \mathbf{M_{qvi}}(p,q) \implies 0 \in p - x + \{0\} \implies x = p, L_p = 1, L_q = 0 < 1.$$

For the parametric VI,

$$0 \in \mathbf{M_{vi}}(p,q) \implies 0 \in p - x + \{0\} \implies x = p, L_p = 1, L_q = 0 < \frac{1}{2}.$$

Hence, Assumptions 4.3 and 4.4 hold. However, analyzing the mapping in (6.3) with $r = 1$,

$$x \in \mathbf{M^r}(p,q) \implies 0 \in p - x + x + \{0\} \implies p = 0, x \in \mathbb{R}^n.$$

Therefore,

$$\mathbf{M^r}(p,q) = \begin{cases} \mathbb{R}^n, \text{ if } p = 0, \\ \emptyset, \text{ if } p \neq 0. \end{cases}$$

Clearly, Assumption 6.1 is not satisfied.

**Example 6.3.** In QVI (1.2), let $K : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be defined by $K(x) = x + \mathbb{R}^n_+$ and $F : \mathbb{R}^n \to \mathbb{R}^n$ be the identity $F(x) = x$.

Any point in the nonnegative orthant $x \in \mathbb{R}^n_+$ is a solution of the QVI, because

$$0 \in F(x) + N_{K(x)}(x) \iff 0 \in x + \mathbb{R}^n_- \iff x \in \mathbb{R}^n_+$$

We can choose a solution to analyze the local assumptions. Take $\bar{x} = (1, \cdots, 1)$. The lack of uniqueness of the QVI solution implies Assumption 4.4 cannot be satisfied; we next consider Assumption 4.3.

Take $(p,q)$ such that $\|p\| \leq \frac{1}{2}, \|q - \bar{x}\| \leq \frac{1}{2}$. We have that

$$0 \in \mathbf{M_{vi}}(p,q) \iff 0 \in p + x + N_{K(q)}(x). \tag{6.4}$$

It is easy to verify that $x = q$ satisfies (6.4). In fact,

$$\begin{aligned} |p_i| \leq \|p\| \leq \frac{1}{2} \quad &\implies \quad p_i \geq -\frac{1}{2}, \\ |q_i - \bar{x}_i| \leq \|q - \bar{x}\| \leq \frac{1}{2} \implies q_i - \bar{x}_i \geq -\frac{1}{2} \quad &\implies \quad q_i \geq \frac{1}{2}, \\ p_i + q_i \geq 0 \implies p + q \in \mathbb{R}^n_+ \quad &\implies \quad 0 \in p + q + \mathbb{R}^n_- = p + q + N_{K(q)}(q). \end{aligned}$$

Thus, by the strong monotonicity of $x \mapsto p + x$, we conclude that $q$ is the unique solution: $q = \mathbf{M_{vi}}(p,q)$. Although we have local single-valuedness of $\mathbf{M_{vi}}$, Assumption 4.3 is not satisfied, because $L_q = 1 > \frac{1}{2}$, contrary to required.

We now examine the mapping in (6.3). We have that

$$x \in \mathbf{M^r}(p,q) \iff 0 \in x + p + r(x - \bar{x}) + N_{K(q)}(x). \tag{6.5}$$

Take $(p,q)$ such that $\|p\| < \frac{1}{4}, \|q - \bar{x}\| < \min\{\frac{1}{4}, \frac{1}{4r}\}$. Taking $x = q$ in (6.5), we have $N_{K(q)}(q) = \mathbb{R}^n_-$. Note that

$$\begin{aligned} \|p + r(q - \bar{x})\| \leq \|p\| + r\|q - \bar{x}\| < \frac{1}{2} \quad &\implies \quad |p_i + r(q_i - \bar{x}_i)| < \frac{1}{2} \implies p_i + r(q_i - \bar{x}_i) > -\frac{1}{2}, \\ \|q - \bar{x}\| < \frac{1}{4} \implies |q_i - \bar{x}_i| < \frac{1}{4} \quad &\implies \quad q_i - \bar{x}_i > -\frac{1}{4} \implies q_i > 1 - \frac{1}{4} = \frac{3}{4}, \\ q_i + p_i + r(q_i - \bar{x}_i) > \frac{3}{4} - \frac{1}{2} > 0 \quad &\implies \quad q + p + r(q - \bar{x}) \in \mathbb{R}^n_+. \end{aligned}$$

Finally,

$$0 \in q + p + r(q - \bar{x}) + \mathbb{R}^n_- = p + q + r(q - \bar{x}) + N_{K(q)}(q) \implies q \in \mathbf{M^r}(p,q).$$

Strong monotonicity of $x \mapsto x + p + r(x - \bar{x})$ implies that the solution is unique. Thus, $q = \mathbf{M^r}(p,q)$, $L'_p = 0, L'_q = 1$ and Assumption 6.1 is satisfied.

We are now ready to state the localization assumption required in [15], given in (6.6) below.

**Proposition 6.4.** *Under Assumption 6.1, there exists an $\varepsilon > 0$ such that for every $(\mathbf{x}, \mathbf{y}) \in \mathcal{W}_\varepsilon$, there exists a pair $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ such that*

$$(\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}), \tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}})) \in \operatorname{gph} \mathbf{T} \cap \mathbf{U}, \tag{6.6}$$

*where $\mathbf{U}$ is the semimonotonicity neighborhood as given in inequality (4.2) and Theorem 4.9.*

*Proof.* Choose $\delta > 0$ such that $\{(\bar{\mathbf{x}}, \mathbf{0})\} + \mathbb{B}_\delta \subset \mathbf{U}$. Take $\delta_1, \delta_2 > 0$ such that $\mathbb{B}_{\delta_1} \subset V'_p$ and $\{\bar{x}\} + \mathbb{B}_{\delta_2} \subset V'_q$. First, take $\varepsilon > 0$ such that

$$\varepsilon \leq \varepsilon_0 = \min\left\{ \delta_1 \left( r\sqrt{\frac{\lambda_x}{2r}} + \sqrt{\frac{r\lambda_y}{2}} \right)^{-1}, \delta_2 \left( \sqrt{\frac{\lambda_x}{2r}} + \frac{1}{r}\sqrt{\frac{r\lambda_y}{2}} \right)^{-1} \right\}. \tag{6.7}$$

Given $(\mathbf{x}, \mathbf{y}) = \left( \begin{bmatrix} x \\ x \end{bmatrix}, \begin{bmatrix} y \\ -y \end{bmatrix} \right) \in \mathcal{W}_\varepsilon$, take $\tilde{y}_1 = y, \tilde{y}_2 = 0, \tilde{x}_2 = x$, and let $\tilde{x}_1$ be the solution of the $\operatorname{VI}(F + rI - rx - y, K(x - \frac{1}{r}y))$. We have that

$$(\tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}}))_2 = \tilde{y}_2 + r(x - \tilde{x}_2) = 0, \tag{6.8}$$

and

$$0 \in F(\tilde{x}_1) + r\tilde{x}_1 - rx - y + N_{K(x - \frac{1}{r}y)}(\tilde{x}_1) \iff y + r(x - \tilde{x}_1) \in F(\tilde{x}_1) + N_{K(x - \frac{1}{r}y)}(\tilde{x}_1). \tag{6.9}$$

Therefore,

$$(\tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}}))_1 \in F((\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}))_1) + N_{K((\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}))_2)}((\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}))_1). \tag{6.10}$$

Conditions (6.8) and (6.10) are equivalent to

$$(\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}), \tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}})) \in \operatorname{gph} \mathbf{T}.$$

By (6.9), we have that

$$0 \in F(\tilde{x}_1) + r(\tilde{x}_1 - \bar{x}) + r(\bar{x} - x) - y + N_{K(x - \frac{1}{r}y)}(\tilde{x}_1) \iff \tilde{x}_1 = \mathbf{M}^{\mathbf{r}}(r(\bar{x} - x) - y, x - \frac{1}{r}y). \tag{6.11}$$

Using (6.2) and (6.7), we conclude that

$$\begin{aligned} \|r(x - \bar{x}) - y\| \leq r\|x - \bar{x}\| + \|y\| \leq \varepsilon \left( r\sqrt{\frac{\lambda_x}{2r}} + \sqrt{\frac{r\lambda_y}{2}} \right) \leq \delta_1 \\ \implies r(x - \bar{x}) - y \in \mathbb{B}_{\delta_1} \subset V'_p, \end{aligned} \tag{6.12}$$

and

$$\begin{aligned} \|(x - \frac{1}{r}y) - \bar{x}\| \leq \|x - \bar{x}\| + \frac{1}{r}\|y\| \leq \varepsilon \left( \sqrt{\frac{\lambda_x}{2r}} + \frac{1}{r}\sqrt{\frac{r\lambda_y}{2}} \right) \leq \delta_2 \\ \implies x - \frac{1}{r}y \in \{\bar{x}\} + \mathbb{B}_{\delta_2} \subset V'_q. \end{aligned} \tag{6.13}$$

Assumption 6.1 applied to (6.11) implies that

$$\|\tilde{x}_1 - \bar{x}\| \leq L'_p\|r(\bar{x} - x) - y\| + L'_q\|x - \frac{1}{r}y - \bar{x}\|.$$

Using the triangle inequality and (6.2) once again, we have

$$\|\tilde{x}_1 - \bar{x}\| \leq (rL'_p + L'_q)\|x - \bar{x}\| + (L'_p + \frac{1}{r}L'_q)\|y\| \leq \varepsilon \left( (rL'_p + L'_q)\sqrt{\frac{\lambda_x}{2r}} + (L'_p + \frac{1}{r}L'_q)\sqrt{\frac{r\lambda_y}{2}} \right).$$

Denote $K = (rL'_p + L'_q)\sqrt{\frac{\lambda_x}{2r}} + (L'_p + \frac{1}{r}L'_q)\sqrt{\frac{r\lambda_y}{2}}$, so that

$$\|\tilde{x}_1 - \bar{x}\| \leq K\varepsilon. \tag{6.14}$$

We also have

$$\|x - \tilde{x}_1\| \leq \|x - \bar{x}\| + \|\bar{x} - \tilde{x}_1\| \leq \left(\sqrt{\frac{\lambda_x}{2r}} + K\right)\varepsilon. \tag{6.15}$$

Note that

$$\begin{aligned}
& \|\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}) - \bar{\mathbf{x}}\|^2 + \|\tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}}) - \mathbf{0}\|^2 \\
= & \left\| \begin{bmatrix} \tilde{x}_1 \\ x - \frac{1}{r}y \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \bar{x} \end{bmatrix} \right\|^2 + \left\| \begin{bmatrix} y + r(x - \tilde{x}_1) \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\|^2 \\
= & \|\tilde{x}_1 - \bar{x}\|^2 + \|x - \bar{x} - \frac{1}{r}y\|^2 + \|y + r(x - \tilde{x}_1)\|^2 \\
\leq & \|\tilde{x}_1 - \bar{x}\|^2 + 2(\|x - \bar{x}\|^2 + \frac{1}{r^2}\|y\|^2) + 2(\|y\|^2 + r^2\|x - \tilde{x}_1\|^2) \\
\leq & (K\varepsilon)^2 + 2\left(\frac{\lambda_x}{2r}\varepsilon^2 + \frac{1}{r^2}\frac{r\lambda_y}{2}\varepsilon^2\right) + 2\left(\frac{r\lambda_y}{2}\varepsilon^2 + r^2\left(K + \sqrt{\frac{\lambda_x}{2r}}\right)^2\varepsilon^2\right) \\
= & \alpha\varepsilon^2,
\end{aligned}$$

where

$$\alpha = K^2 + \frac{\lambda_x}{r} + \frac{\lambda_y}{r} + r\lambda_y + 2r^2\left(K + \sqrt{\frac{\lambda_x}{2r}}\right)^2.$$

Choosing $\varepsilon \leq \min\{\varepsilon_0, \frac{\delta}{\sqrt{\alpha}}\}$, we obtain that

$$\|(\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}), \tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}})) - (\bar{\mathbf{x}}, \mathbf{0}))\| \leq \sqrt{\alpha}\varepsilon \leq \delta,$$

implying that

$$(\tilde{\mathbf{x}} + \frac{1}{r}(\mathbf{y} - \tilde{\mathbf{y}}), \tilde{\mathbf{y}} + r(\mathbf{x} - \tilde{\mathbf{x}})) \in \mathbf{U}.$$

□

Our final result, stating local convergence, is shown for monotone $F$, so that each VI subproblem has unique solution. This follows from Theorem 2.3.3-(b) in [19], since the proximal term makes the operator strongly monotone. Monotonicity of $F$ is used only to ensure uniqueness of the iterates. In the absence of that property, the same local argument would apply to any sequence generated, by selecting a localized solution at each step, whose existence is guaranteed by Proposition 6.4.

**Theorem 6.5** (Local convergence of Algorithm 2)**.** *Assume that $\bar{x}$ is a solution of QVI 1.2, $F$ is monotone, Assumption 6.1 holds, and the parameters in Algorithm 2 satisfy* (5.1)*, where $\mu, \mu^{\perp} > 0$ with $\mu\mu^{\perp} < 1$ are given by Theorem 4.9. Then the following statements hold whenever the starting point $(x^0, y^0)$ is close enough to $(\bar{x}, 0)$:*

- *(i) If Assumption 4.3 holds, then either a solution of QVI 1.1 is reached in a finite number of iterations, or the sequence $\{x^k\}$ generated by Algorithm 2 is bounded and its every accumulation point is a solution of QVI 1.1. In particular, if $\bar{x}$ is the unique solution of QVI 1.1 in the neighborhood under consideration, then the sequence converges to $\bar{x}$.*
- *(ii) If Assumption 4.4 holds, then either the solution $\bar{x}$ is reached in a finite number of iterations, or the generated sequence $\{x^k\}$ converges to the solution of QVI (unique in this case).*

*Proof.* Assumption III in [15] states the conditions required for local convergence of the progressive decoupling+ algorithm, divided into four items: A1, A2, A3, A4. We proceed to show that these hold in our case.

Theorem 4.9 establishes the semimonotonicity property, which is precisely the requirement in A1.

Item A2 requires outer semicontinuity of $\mathbf{T}$, which holds on the entire domain, as shown in the proof of Theorem 5.2.

Item A3 requires the parameters to be chosen within nonempty sets specified above, with $\mu\mu^{\perp} < 1$. The latter condition is ensured by Theorem 4.9.

Item A4 is precisely (6.6), which holds in our case by Proposition 6.4.

Because of item A4, a localized solution exists at each iteration of Algorithm 2. Since monotonicity of $F$ ensures uniqueness of the VI solution at each step, the iterates are uniquely defined and remain in the semimonotonicity neighborhood.

By item (iii) of Theorem 4.9 in [15], whenever the initial point $(x^0, y^0)$ is sufficiently close to $(\bar{x}, 0)$, equivalently, whenever $(\mathbf{x}^0, \mathbf{y}^0)$ is sufficiently close to $(\bar{\mathbf{x}}, \mathbf{0})$, the sequence $\{x^k\}$ generated by Algorithm 2 either reaches a solution in a finite number of iterations or it is bounded and all its accumulation points are solutions of QVI 1.1. Then if $\bar{x}$ is the unique solution of QVI 1.1 in the neighborhood under consideration, the sequence converges to $\bar{x}$. The case of Assumption 4.4 falls into this local uniqueness setting. $\square$

# 7 Numerical Results

To assess computational performance of our proposal, we apply Algorithm 1 to solve Walrasian equilibrium problems formulated as GNEPs [1]. A GNEP can be solved combining the optimality conditions of agents' problems, which leads to a QVI, as in [17, 18]. We compare our method with a direct solution using the Extended Mathematical Programming (EMP) module in the software GAMS [20], described in [27]. The latter method uses safeguarded semismooth Newton techniques to solve the mixed complementarity reformulation of GNEP, using the PATH solver [10]. We also benchmark Algorithm 1 against the recently proposed QVI Dantzig-Wolfe decomposition method [23].

Before solving the large-scale Walrasian equilibrium problems, we tune the proximal parameter on a battery of QVI problems available in the literature. These experiments were conducted to understand the sensitivity of the method with respect to the proximal parameter $r > 0$, to gain some intuition on what values are suitable for computational purposes. To this end, we apply the method to the academic problems from [16] and to other Walrasian equilibrium problems from [9].

All tests were conducted on a computer running Ubuntu 22.04 with an AMD Ryzen Threadripper 1950X processor featuring 16 cores (32 threads) and 64 GB of RAM. The codes are available at `https://github.com/ManoelJardim/ProgQVI/tree/main/ProgQVI`. We used GAMS with its default parameter settings. All codes were written in Python, with explicit calls to GAMS and other necessary packages when required.

## 7.1 Proximal Parameter Analysis

In [9], numerical examples of the Arrow-Debreu model with utility functions based on the constant elasticity of substitution are presented. We simulated Examples 3.7, 3.8, and 3.9 from that paper. The remaining QVI problems solved in this section are described in [16], where the authors propose a benchmark library of QVI problems called QVILIB. Since the realistic Walrasian problems are already part of our test set, we selected the academic problems from QVILIB.

The experiment includes all 29 academic problems from QVILIB and the 3 examples from [9], amounting to 32 problems.

For each problem, we tested the proximal parameter for the fixed values $r \in \{0.001, 1, 100\}$, and monotonically increasing parameters given by $r_k = \min(0.001 \times 2^k, 10^6)$, starting at 0.001 and bounded above by $10^6$. The initial point was set to zero, $x^0 = y^0 = 0_{\mathbb{R}^n}$, except for four QVILIB instances for which $0_{\mathbb{R}^n}$ is already a solution. In these cases, we set the initial point to $x^0 = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ and $y^0 = 0_{\mathbb{R}^n}$. We counted the number of iterations until the stopping criterion measure fell below $10^{-3}$. The performance profile in Figure 1 indicates the sensitivity of

Algorithm 1 to the choice of the proximal parameter. The performance factor was defined in terms of the number of iterations, which was set to infinity for cases in which the algorithm diverged.

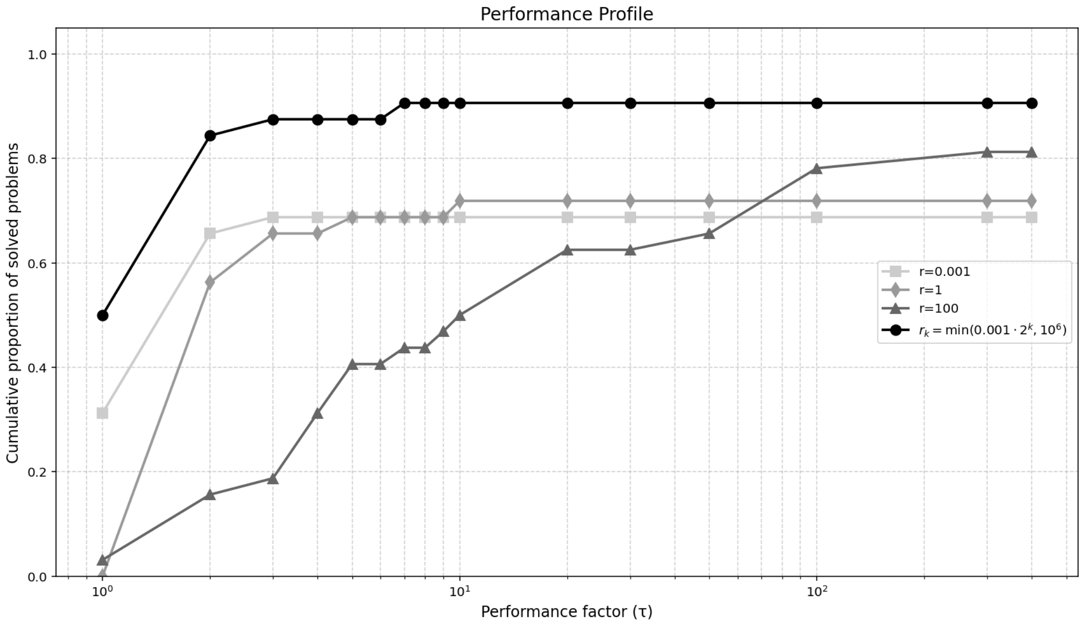


Figure 1: Performance profile comparing different proximal parameters.

Convergence was observed for almost all problems for at least one choice of the proximal parameter or by using the increasing parameter. Only the Box1B problem from QVILIB did not converge for any of the parameters tested. This problem is known to be difficult and has been reported in the literature [24, 25], with most methods failing to converge except for exact penalization techniques in [25]. Algorithm 1 is based on solving a variational inequality (VI) subproblem at each iteration; therefore, its convergence depends on the robustness of the subproblem solver. This explains the divergence observed for the Box1B problem, which poses challenges to VI solvers in its subproblems.

Smaller values of the proximal parameter led to faster convergence. On the other hand, some problems required larger values of $r$ to converge, which could be mitigated by using a variable proximal parameter. The increasing parameter strategy achieved convergence in fewer iterations in most cases, combining the faster convergence of small proximal parameter values with the stability obtained for larger values.

## 7.2 Walrasian Equilibrium Problem

We closely follow the Walrasian problem formulation in [23]. There are $C$ consumers, $G$ goods, one firm and one regulator doing the market clearance, totaling $C + 2$ players with $G$ goods.

### 7.2.1 VI subproblems

For $i = 1, ..., C$, denote $x^i \in \mathbb{R}^G$ the goods to be bought by consumer $i$. To avoid confusion, in this section, iteration indices will be written in parentheses, such as $x^{(k)}$ or $x^{i,(k)}$. Each consumer has a concave utility function $\mathcal{U}^i(x^i)$ and a budget constraint defined by its endowment $\mathcal{E}^i \in \mathbb{R}^G$:

$$\max_{x^i \geq 0} \left\{ \mathcal{U}^i(x^i) : \langle p, x^i \rangle \leq \langle p, \mathcal{E}^i \rangle \right\}. \tag{7.1}$$

The firm produces a quantity of goods, denoted by $x^{C+1} \in \mathbb{R}^G$, to maximize its profit under a production constraint, but it could be absent in a purely exchange economy, having the problem

$$\max_{x^{C+1}\geq 0}\left\{\langle p, x^{C+1}\rangle : \sum_{j=1}^{G}(x_j^{C+1})^2 \leq M\right\}. \tag{7.2}$$

The price of each good $p^i \in \mathbb{R}^G$ is set by the market player to achieve market clearance, when there is an equilibrium between demand and supply, equivalent to Walras's law,

$$\max_{p\geq 0}\left\{\left\langle p, \sum_{i=1}^{C}(x^i - \mathcal{E}^i) - x^{C+1}\right\rangle : \sum_{j=1}^{G} p_j = 1\right\}. \tag{7.3}$$

Labeling the market player with $C+2$ so that $p = x^{C+2}$, $x = \begin{bmatrix} x^1 \\ \vdots \\ x^{C+2}\end{bmatrix}$, the GNEP can be solved by the following QVI$(F,K)$:

$$F(x) = \Big(-\nabla_{x^1}\mathcal{U}^1(x^1), \dots, -\nabla_{x^C}\mathcal{U}^C(x^C), -x^{C+2}, \sum_{i=1}^{C}(\mathcal{E}^i - x^i) + x^{C+1}\Big), \tag{7.4}$$

$$K(x) = \{z \in \mathcal{D} \subset \mathbb{R}^{G(C+2)} : \sum_{j=1}^{G} x_j^{C+2}(z_j^i - \mathcal{E}_j^i) \leq 0, \text{ for } 1 \leq i \leq C\}, \tag{7.5}$$

$$\mathcal{D} = \{z \in \mathbb{R}^{G(C+2)} : z \geq 0, \sum_{j=1}^{G}(z_j^{C+1})^2 \leq M, \sum_{j=1}^{G} z_j^{C+2} = 1\}. \tag{7.6}$$

The main cost of the iteration in our algorithm consists in solving a VI. This VI can be decoupled employing the structure of the convex set $K^{(k)}$. We next write explicitly the iterations VI for the mappings (7.4),(7.5), simplifying the notation by renaming the vectors for the calculations: $x = \hat{x}^{(k)}, d = y^{(k)} + rx^{(k)}$ and $q = x^{(k)} - \frac{1}{r}y^{(k)}$. This gives the following.

$$\text{Find } x \in K^{(k)} \text{ s.t. } \langle F(x) + rx - d, z - x\rangle \geq 0, \ \ \forall z \in K^{(k)},$$

which means

$$\begin{aligned}
\text{Find } x \in K^{(k)} \text{ s.t.} \qquad & \sum_{i=1}^{C}\langle -\nabla_{x^i}\mathcal{U}^i(x^i) + rx^i - d^i, z^i - x^i\rangle \\
+ \quad & \langle -x^{C+2} + rx^{C+1} - d^{C+1}, z^{C+1} - x^{C+1}\rangle \\
+ \quad & \langle \sum_{i=1}^{C}(\mathcal{E}^i - x^i) + x^{C+1} + rx^{C+2} - d^{C+2}, z^{C+2} - x^{C+2}\rangle \geq 0, \\
& \forall z \in K^{(k)}.
\end{aligned} \tag{7.7}$$

Note that

$$K^{(k)} = \{z \in \mathcal{D} : \sum_{j=1}^{G} q_j^{C+2}(z_j^i - \mathcal{E}_j^i) \leq 0, \text{for } 1 \leq i \leq C\} = \prod_{i=1}^{C} D^i \times D^{C+1} \times D^{C+2},$$

$$\text{where } D^i = \{z^i \in \mathbb{R}_+^G : \sum_{j=1}^{G} q_j^{C+2}(z_j^i - \mathcal{E}_j^i) \leq 0\},$$

$$D^{C+1} = \{z^{C+1} \in \mathbb{R}_+^G : \sum_{j=1}^{G}(z_j^{C+1})^2 \leq M\}, D^{C+2} = \{z^{C+2} \in \mathbb{R}_+^G : \sum_{j=1}^{G} z_j^{C+2} = 1\}.$$

For each $i \in \{1, ..., C\}$, take $z \in K^{(k)}$, with $z^m = x^m$ when $m \neq i, z^i \in D^i$ arbitrary, $z^{C+1} = x^{C+1}$ and $z^{C+2} = x^{C+2}$. Applying this in (7.7), we get VIs in a small-dimensional space $\mathbb{R}^G$:

$$\text{Find } x^i \in D^i \text{ s.t. } \langle -\nabla_{x^i}\mathcal{U}^i(x^i) + rx^i - d^i, z^i - x^i\rangle \geq 0, \ \forall z^i \in D^i. \tag{7.8}$$

Taking $z \in K^{(k)}$, with $z^i = x^i$ for $1 \le i \le C, z^{C+1} \in D^{C+1}, z^{C+2} \in D^{C+2}$, applying these in (7.7), we get the VI

$$\text{Find } \begin{bmatrix} x^{C+1} \\ x^{C+2} \end{bmatrix} \in D^{C+1} \times D^{C+2} \quad \text{s.t.}$$

$$\left\langle \begin{bmatrix} -x^{C+2} + rx^{C+1} - d^{C+1} \\ \sum_{i=1}^{C} (\mathcal{E}^i - x^i) + x^{C+1} + rx^{C+2} - d^{C+2} \end{bmatrix}, \begin{bmatrix} z^{C+1} \\ z^{C+2} \end{bmatrix} - \begin{bmatrix} x^{C+1} \\ x^{C+2} \end{bmatrix} \right\rangle \ge 0,$$

$$\forall \begin{bmatrix} z^{C+1} \\ z^{C+2} \end{bmatrix} \in D^{C+1} \times D^{C+2}. \qquad (7.9)$$

We still have the variables $x^i, i \in \{1, ..., C\}$, in VI (7.9), but we can first solve the VIs in (7.8) to obtain these vectors $x^i$. Crucially, due to their structure, each of these VIs can be solved separately as a convex optimization problem. Once these vectors are determined, they are no longer variables in (7.9), which then becomes a VI on $\mathbb{R}^{2G}$. This decoupling allows us to solve at each iteration $C$ separate $G$-dimensional VIs (as convex optimization problems!), followed by a single $2G$-dimensional VI, rather than tackling the large $G(C+2)$-dimensional VI. This feature appears to be a great gain for large-scale problems.

### 7.2.2 Computational Tests

Our experiments compare the following three approaches:

1. Direct Method (DIRECT): using the EMP module from GAMS [27], based on PATH solver. The parameters are used in default mode of GAMS;
2. Dantzig-Wolfe decomposition (DW): The DW decomposition for QVI proposed in [23]. The VI subproblems are solved using convex programming, with CVXOPT package in Python for the quadratic case, and QVI master problems are solved using EMP. The initial point takes prices equal to $1/G$ and zero for the other values, while the stopping criterion requires to have GAP $\ge -0.01$;
3. Progressive decoupling of QVI (ProQVI): Algorithm 1. Each VI decoupled in (7.8) solved using convex programming, with CVXOPT package in Python for the quadratic case. The nonsymmetric VI in (7.9) is solved using GAMS. Based on the sensitivity analysis in section 7.1, and opting for a fixed proximal parameter, we use $r = 0.001$. The initial point sets the prices equal to $1/G$, and assigns zero to the other values of $x^0$ and to $y^0$. The stopping criterion is based on tol $= 0.001$.

The data is generated randomly as done in [23]. The utility functions are concave and quadratic $\mathcal{U}^i(x^i) = -\frac{1}{2}\langle x^i, \mathcal{R}^i x^i \rangle + \langle b^i, x^i \rangle$, for $i = 1, ..., C$, where $b^i \in \mathbb{R}^G$ has elements uniformly distributed in $[0, 10]$, and $\mathcal{R}^i$ is a $G \times G$ positive semidefinite matrix generated by $\frac{10}{\|A^{iT} A^i\|_\infty} A^{iT} A^i$, with $A^i \in \mathbb{R}^{G \times G}$ having elements uniformly distributed between $[-1, 1]$, so that the elements of $\mathcal{R}^i$ are between $[-10, 10]$, regardless of the size of $G$. The endowments are also obtained randomly with uniform distribution between $[0, 10]$ and the firm capacity is proportional to $CG$, maintained at levels sufficient to meet market demand; however, this parameter has minor influence on the numerical results.

The problem can increase in dimension by having more consumers $C$ or having more goods $G$, or even increasing both simultaneously. While we have $G(C+2)$ variables, we have $C$ main constraints in the definition of $K(x)$ in (7.5), besides the constraints in the definition of the set $\mathcal{D}$. For our analysis, we increase $G$ and $C$ simultaneously, in three different benchmarks, based on the relationships $C/G \in \{0.5, 1, 2\}$, allowing to evaluate big problems with different structure concerning the dimensions of the variables and constraints.

The results in Table 1 show that both decomposition methods behave very well for large scale problems, surpassing the direct approach for these instances. For small problems, the direct approach is advantageous, since the problem is solved almost instantaneously. This was already observed when testing the DW approach [23]. For the ProQVI, we observe a very good performance, even surpassing the DW decomposition for several cases. The reason is that after decoupling, all the subproblems in ProQVI decomposition are simple VIs (some even convex optimization problems), not needing any QVI solution (like DW decomposition). Particularly, for problems with more constraints, the ProQVI approach

Table 1: CPU time for solvers DIRECT, DW and ProQVI, and number of DW and ProgQVI iterations with 600 instances of the Walrasian equilibrium problem. In each row, the solver in bold face is the one having the lowest mean time of execution. The three cases marked with $*$ had 1 divergence in 20 simulations, while the two marked with # had 2 divergences between 20 simulations.

| $(C, G, n)$ | DIRECT time (s) | | DW time (s) | | PROG time (s) | | Iterations (mean) | |
|---|---|---|---|---|---|---|---|---|
| | mean | max | mean | max | mean | max | DW | PROG |
| (20, 20, 440) | **0.2** | 0.2 | 2.3 | 4.1 | 2.1 | 4.5 | 16.2 | 33.6 |
| (40, 40, 1,680) | **0.8** | 1.1 | 3.8 | 7.3 | 2.9* | 5.9 | 18.8 | 39.5 |
| (60, 60, 3,720) | 6.4 | 74.1 | 4.3 | 8.4 | **4.2** | 6.4 | 18.2 | 58.5 |
| (80, 80, 6,560) | 23.5 | 220.4 | **4.8** | 7.0 | 6.2 | 8.6 | 15.9 | 80.6 |
| (100, 100, 10,200) | 256.9 | 1,710.5 | 4.7 | 8.7 | **4.1** | 10.8 | 13.0 | 54.0 |
| (120, 120, 14,640) | 702.9 | 5,887.3 | 5.8 | 9.7 | **4.0** | 4.7 | 12.7 | 48.8 |
| (140, 140, 19,880) | 583.8 | 3,855.4 | 5.8 | 9.6 | **4.5** | 5.7 | 10.5 | 54.4 |
| (160, 160, 25,920) | 1,660.3 | 13,803.8 | 6.5 | 10.9 | **5.8** | 6.6 | 9.1 | 60.5 |
| (180, 180, 32,760) | 1,867.2 | 4,568.6 | 7.4 | 13.9 | **6.7** | 9.8 | 8.2 | 60.5 |
| (200, 200, 40,400) | 1,844.7 | 10,596.5 | **7.7** | 17.9 | 7.8 | 9.1 | 6.8 | 73.1 |
| (10, 20, 240) | **0.1** | 0.1 | 1.3 | 3.0 | 1.9 | 2.2 | 10.4 | 29 |
| (20, 40, 880) | **0.4** | 0.4 | 1.8 | 2.8 | 2.3 | 3.0 | 13.0 | 35.7 |
| (30, 60, 1,920) | **1.2** | 1.3 | 2.3 | 3.9 | 3.9 | 5.1 | 13.8 | 54.9 |
| (40, 80, 3,360) | 2.9 | 3.1 | **2.3** | 4.6 | 4.6 | 6.5 | 11.9 | 60.4 |
| (50, 100, 5,200) | 6.0 | 7.6 | **2.1** | 4.5 | 5.3 | 8.2 | 10.3 | 67.6 |
| (60, 120, 7,440) | 60.8 | 915.8 | **2.5** | 3.6 | 4.8 | 9.9 | 9.6 | 54.9 |
| (70, 140, 10,080) | 33.9 | 65.1 | **2.5** | 4.5 | 3.5 | 4.1 | 7.5 | 41.7 |
| (80, 160, 13,120) | 139.1 | 1,731.9 | **3.6** | 6.0 | 4.6 | 5.3 | 8.5 | 49.4 |
| (90, 180, 16,560) | 77.5 | 213.7 | **3.8** | 6.0 | 5.4 | 6.3 | 7.2 | 54 |
| (100, 200, 20,400) | 88.6 | 333.1 | **3.1** | 6.0 | 5.5 | 7.0 | 5.3 | 55.3 |
| (20, 10, 220) | **0.1** | 0.1 | 3.1 | 5.1 | 1.8* | 2.0 | 19.8 | 29.5 |
| (40, 20, 840) | **0.3** | 0.6 | 4.8 | 8.2 | 2.0* | 2.5 | 24.4 | 32.9 |
| (60, 30, 1,860) | **0.8** | 1.3 | 8.3 | 16.1 | 2.6 | 3.8 | 28.9 | 40.0 |
| (80, 40, 3,280) | **2.6** | 14.2 | 12.2 | 24.0 | 3.4 | 5.1 | 31.3 | 50.7 |
| (100, 50, 5,100) | 15.8 | 57.5 | 15.0 | 41.9 | **4.5** | 6.4 | 30.5 | 66.8 |
| (120, 60, 7,320) | 69.9 | 195.3 | 15.6 | 33.2 | **5.1** | 7.6 | 29.0 | 72.9 |
| (140, 70, 9,940) | 169.3 | 345.9 | 17.4 | 87.8 | **4.5**# | 8.5 | 27.0 | 62.8 |
| (160, 80, 12,960) | 343.2 | 868.5 | 16.5 | 49.3 | **3.3**# | 7.3 | 26.0 | 53.7 |
| (180, 90, 16,380) | 699.8 | 2,741.0 | 14.4 | 25.3 | **3.7** | 7.8 | 22.0 | 52.0 |
| (200, 100, 20,200) | 1,122.3 | 3,836.6 | 16.3 | 45.8 | **4.1** | 4.8 | 21.5 | 54.3 |

is naturally superior, since the DW technique needs to solve its QVI master subproblems. However, for problems with fewer constraints, DW was superior to ProQVI, although the execution time was similar. In conclusion, both ProQVI and DW decomposition are useful for large problems, with the number of constraints in the problem determining which one of the two approaches is superior.

The graphs in Figure 2 show the trend for the simulations, considering the statistical time, since we have solved 20 problems for each case, using the same starting points. As we can see, both decompositions present a small variation of time for the problems, and the ProQVI has the smallest dispersion, except for a few isolated instances which presented divergence (1.17% of the cases for $r = 0.001$).

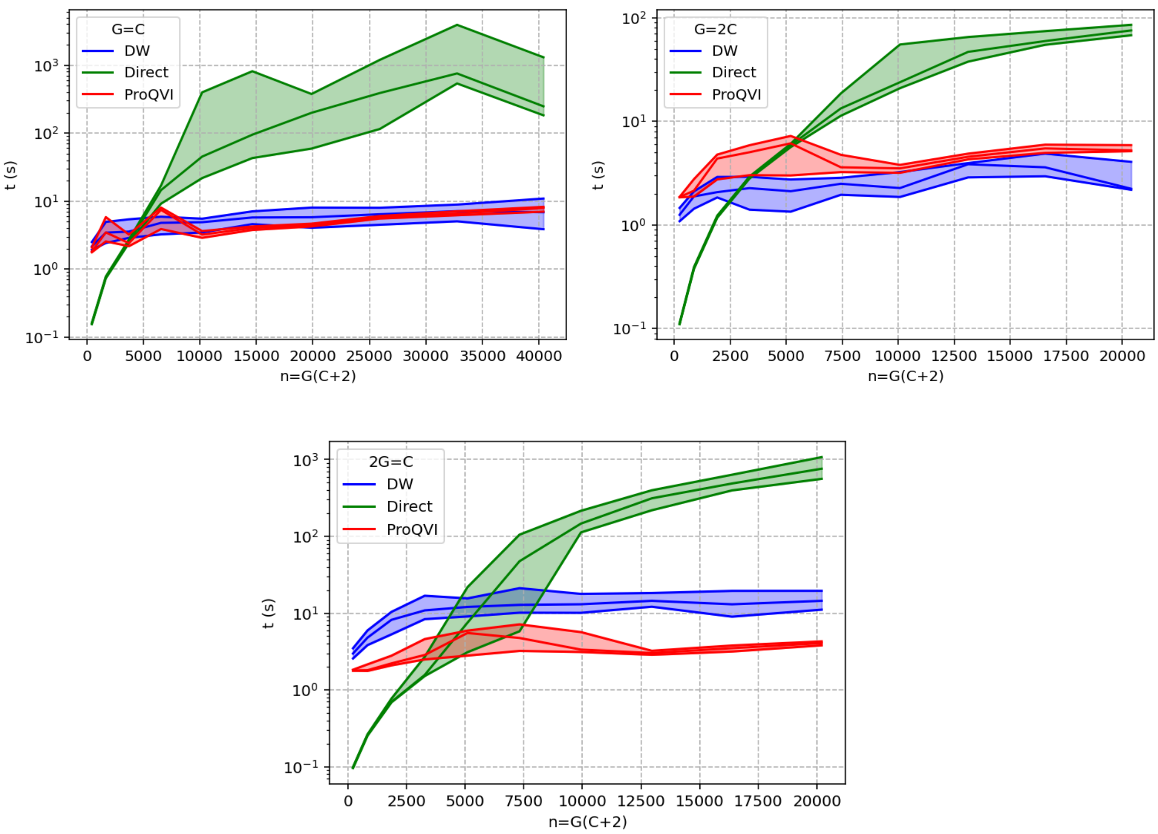


Figure 2: Time statistics (median and 25%-75% quantiles) for DIRECT, DW and ProQVI, increasing $n$ with $G = C$, $G = 2C$, and $2G = C$.

# Declarations

**Conflicts of Interest.** The authors declare that they have no conflict of interest of any kind related to the manuscript.
**Data Availability Statement.** Data sharing is not applicable to this article.

# Acknowledgements

The second author is supported by CNPq Grant 307509/2023-0. The third author is supported in part by CNPq Grant 306775/2023-9 and by FAPERJ Grant E-26/200.165/2026.

The authors thank Professor Paulo Silva from UNICAMP for providing access to the computational resources necessary for conducting our numerical experiments.

## References


[1] Kenneth J. Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3):265, July 1954.

[2] D. Aussel and J. Cotrina. Quasimonotone quasivariational inequalities: Existence results and applications. *Journal of Optimization Theory and Applications*, 158(3):637–652, February 2013.

[3] D. Aussel, M. B. Donato, M. Milasi, and A. Sultana. Existence results for quasi-variational inequalities with applications to Radner equilibrium problems: Resolution through variational inequalities. *Set-Valued and Variational Analysis*, 29(4):931–948, September 2021.

[4] Assis Azevedo, Fernando Miranda, and Lisa Santos. Variational and quasivariational inequalities with first order constraints. *Journal of Mathematical Analysis and Applications*, 397(2):738–756, January 2013.

[5] A. Bensoussan, M. Goursat, and J.-L. Lions. Contrôle impulsionnel et inéquations quasi-variationnelles stationnaires. *C. R. Acad. Sci. Paris Sér. A*, 276:1279–1284, 1973.

[6] A. Bensoussan and J.-L. Lions. Nouvelle formulation de problèmes de contrôle impulsionnel et applications. *C. R. Acad. Sci. Paris Sér. A*, 276:1189–1192, 1973.

[7] A. Bensoussan and J.-L. Lions. Nouvelles méthodes en contrôle impulsionnel. *Appl. Math. Optim.*, 1:289–312, 1975.

[8] D. Chan and J. S. Pang. The generalized quasi-variational inequality problem. *Mathematics of Operations Research*, 7(2):211–222, May 1982.

[9] Julio Deride, Alejandro Jofré, and Roger J-B Wets. Solving deterministic and stochastic equilibrium problems via augmented Walrasian. *Computational Economics*, 53(1):315–342, September 2017.

[10] Steven P. Dirkse and Michael C. Ferris. The PATH solver: a nommonotone stabilization scheme for mixed complementarity problems. *Optimization Methods and Software*, 5(2):123–156, January 1995.

[11] Maria Bernadette Donato, Monica Milasi, and Carmela Vitanza. An existence result of a quasi-variational inequality associated to an equilibrium problem. *Journal of Global Optimization*, 40(1–3):87–97, July 2007.

[12] A. L. Dontchev and R. T. Rockafellar. Regularity and conditioning of solution mappings in variational analysis. *Set-Valued Analysis*, 12(1-2):79–109, March 2004.

[13] Asen L. Dontchev. *Lectures on Variational Analysis.* Springer International Publishing, 2021.

[14] Axel Dreves. Uniqueness for quasi-variational inequalities. *Set-Valued and Variational Analysis*, 24(2):285–297, August 2015.

[15] Brecht Evens, Puya Latafat, and Panagiotis Patrinos. Spingarn's method and progressive decoupling beyond elicitable monotonicity. *arXiv*, abs/2504.00836, 2025.

[16] F. Facchinei, C. Kanzow, and S. Sagratella. QVILIB: A library of quasi-variational inequality test problems. 2012.

[17] Francisco Facchinei and Christian Kanzow. Generalized Nash equilibrium problems. *4OR*, 5(3):173–210, September 2007.

[18] Francisco Facchinei, Christian Kanzow, and Simone Sagratella. Solving quasi-variational inequalities via their KKT conditions. *Mathematical Programming*, 144(1–2):369–412, February 2013.

[19] Francisco Facchinei and Jong-Shi Pang. *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer Series in Operations Research and Financial Engineering. Springer, New York, 2004.

[20] GAMS. *General Algebraic Modeling System Release 40.3.0.* Fairfax, VA, USA, 2022. Release Date: Sep 16, 2022. Available for download at `https://www.gams.com/download/`.

[21] Patrick T. Harker. Generalized Nash games and quasi-variational inequalities. *European Journal of Operational Research*, 54(1):81–94, September 1991.

[22] Michael Hintermüller and Carlos N. Rautenberg. Parabolic quasi-variational inequalities with gradient-type constraints. *SIAM Journal on Optimization*, 23(4):2090–2123, 2013.

[23] Manoel Jardim, Claudia Sagastizábal, and Mikhail Solodov. A Dantzig-Wolfe decomposition method for quasi-variational inequalities. *arXiv*, abs/2505.08108, 2026. arXiv:2505.08108 [math.OC].

[24] Christian Kanzow. On the multiplier-penalty-approach for quasi-variational inequalities. *Mathematical Programming*, 160(1–2):33–63, January 2016.

[25] Christian Kanzow and Daniel Steck. Augmented Lagrangian and exact penalty methods for quasi-variational inequalities. *Computational Optimization and Applications*, 69(3):801–824, November 2017.

[26] Christian Kanzow and Daniel Steck. Quasi-variational inequalities in Banach spaces: Theory and augmented Lagrangian methods. *SIAM Journal on Optimization*, 29(4):3174–3200, January 2019.

[27] Youngdae Kim and Michael C. Ferris. Solving equilibrium problems using extended mathematical programming. *Mathematical Programming Computation*, 11(3):457–501, March 2019.

[28] Alexander S. Kravchuk and Pekka J. Neittaanmäki. *Variational and Quasi-Variational Inequalities in Mechanics*. Springer, Netherlands, 2007.

[29] M. Kunze and J. F. Rodrigues. An elliptic quasi-variational inequality with gradient constraints and some of its applications. *Mathematical Methods in the Applied Sciences*, 23(10):897–908, 2000.

[30] N. Mijajlović and M. Jaćimović. A proximal method for solving quasi-variational inequalities. *Computational Mathematics and Mathematical Physics*, 55(12):1981–1985, December 2015.

[31] Boris Mordukhovich. Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis. *Transactions of the American Mathematical Society*, 343(2):609, 1994.

[32] Boris S. Mordukhovich and Jiří V. Outrata. Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. *SIAM Journal on Optimization*, 18(2):389–412, January 2007.

[33] Yurii Nesterov and Laura Scrimali. Solving strongly monotone variational and quasi-variational inequalities. *Discrete Continuous Dynamical Systems*, 31:1383–1396, 2011.

[34] J. V. Outrata. Solution behaviour for parameter-dependent quasi-variational inequalities. *RAIRO - Operations Research*, 30(4):399–415, 1996.

[35] Michael Patriksson and R. Tyrrell Rockafellar. Variational geometry and equilibrium. In Patrizia Daniele, Franco Giannessi, and Antonino Maugeri, editors, *Equilibrium Problems and Variational Models*, pages 347–368. Springer US, Boston, MA, 2003.

[36] Stephen M. Robinson. Strongly regular generalized equations. *Mathematics of Operations Research*, 5(1):43–62, February 1980.

[37] R. Tyrrell Rockafellar. Progressive decoupling of linkages in optimization and variational inequalities with elicitable convexity or monotonicity. *Set-Valued and Variational Analysis*, 27(4):863–893, October 2018.

[38] L. Scrimali. Quasi-variational inequalities in transportation networks. *Mathematical Models and Methods in Applied Sciences*, 14(10):1541–1560, October 2004.

[39] Jonathan E. Spingarn. Partial inverse of a monotone operator. *Applied Mathematics and Optimization*, 10(1):247–265, June 1983.