A POLICY ITERATION ALGORITHM FOR NONZERO-SUM STOCHASTIC IMPULSE GAMES

René Aïd[1], Francisco Bernal[2], Mohamed Mnif[3], Diego Zabaljauregui[4] and Jorge P. Zubelli[5]

Abstract. This work presents a novel policy iteration algorithm to tackle nonzero-sum stochastic impulse games arising naturally in many applications. Despite the obvious impact of solving such problems, there are no suitable numerical methods available, to the best of our knowledge. Our method relies on the recently introduced characterisation of the value functions and Nash equilibrium via a system of quasi-variational inequalities. While our algorithm is heuristic and we do not provide a convergence analysis, numerical tests show that it performs convincingly in a wide range of situations, including the only analytically solvable example available in the literature at the time of writing.

Résumé. Ce travail présente un nouvel algorithme d'itération sur les politiques pour approximer numériquement les fonctions valeurs d'un problème de jeux impulsionnels stochastiques à somme non nulle. Ces problèmes apparaissent naturellement de nombreuses situations économiques de concurrence entre acteurs. A notre connaissance, malgré l'intérêt pratique de solutions numériques à de tels problèmes, il n'existe pas d'algorithmes appropriées. Notre méthode repose sur la caractérisation récemment introduite des fonctions valeur et de l'équilibre de Nash par un système d'inégalités quasi-variationnelles. Bien que l'on ne fournisse pas d'analyse de convergence, des tests numériques effectués dans un large éventail de situations illustrent l'efficacité de notre algorithme. Enfin, nous montrons qu'il converge à la solution dans le seul cas connu de solution analytique.

INTRODUCTION

Stochastic impulse games (SIGs) are at the intersection between differential game theory and stochastic impulse control. In the case of one sole player, they reduce to stochastic impulse optimisation problems where an agent seeks to control an *underlying* (or *state variable*)—otherwise governed by a stochastic differential equation—in order to maximize the expected value of some target functional. One way in which she can influence the underlying is this: whenever it leaves a *continuation region* (i.e. at her *intervention times*), she suddenly shifts it back somewhere into the region (by providing an *impulse*). Together, the continuation region

[1] Université Paris Dauphine, PSL Research University, LEDA.

[2] CMAP, Ecole Polytechnique, and Finance for Energy Market Research Centre (FiME).

[3] University of Tunis El Manar, ENIT-LAMSIN.

[4] Corresponding Author. Department of Statistics, London School of Economics and Political Science.
E-mail: `d.zabaljauregui@lse.ac.uk`

[5] IMPA.

and impulses define what we will call a *strategy*. Thus, the control-theoretical problem typically boils down to determining an optimal strategy as a function of the current value of the underlying. The purpose of the analysis is to apply that optimal strategy to any specific realization of the underlying. When such a strategy is followed, the expected value of the target functional—the largest possible—is called the *value function*. This problem is a classical one and it is well understood by now [10]. In particular, a rigorous framework has been put in place—combining Howard's algorithm [3–5, 9] with viscosity-capturing finite difference schemes—which allows for robust numerical approximations [7, 8, 10].

A natural extension are two-player SIGs. In such games, two players seek to control the underlying with different, and often opposed, aims. Let us give two concrete examples, due to [1].

- Two central banks competing to influence the exchange rate between their respective currencies, both seeking to devalue their own one. In this application, the exchange rate is modelled as a stochastic process, and either central bank intervenes when it deems its currency too strong. Each bank's strategy is made up of the ensuing devaluation and the threshold exchange rate triggering it. Both quantities are to be optimally determined in advance, by solving the game.
- A wholesale producer of energy and a big client of hers who resells it. Due to the high consumption of the client, both players are capable of affecting the wholesale price of energy, with naturally opposed targets (a high price versus a low one, respectively). Each player incurs costs when modifying the price while possibly benefiting from state compensation upon adverse price movements led by the opponent. In order to avoid superfluous costs, both players need a strategy, based on the current wholesale price, determining the optimal triggering threshold and the extent of the price shift.

Contrary to one-player SIGs, the value functions in two-player SIGs (one for each player) cannot be naively defined by maximisation. Instead, optimality can naturally be characterized via the notion of *Nash equilibrium*: intuitively, the pair of value functions displays the best expected outcome in the sense that if a player changes her strategy while her opponent does not, then the former can only be worse off. A pair of strategies at which these values are attained is then called a Nash equilibrium. Because of this, two-player SIGs are distinctly different (and more complex) from one-player SIGs, and the theoretical/numerical frameworks for the latter are not suitable for the former. This is further accentuated in the so-called *nonzero-sum* SIGs (NZSSIGs), roughly meaning that the losses (resp. gains) undergone by one player may not exactly translate into gains (resp. losses) for the opponent. (For instance, the previously described examples are best modelled under this very flexible framework.) NZSSIGs are in contraposition with the simpler and more particular *zero-sum* SIGs, in which both value functions add up to zero—effectively reducing the problem to finding one. (This allows for more natural extensions of the theoretical and numerical tools of the one-player case [2, 6].)

Due to the greater complication involved in the analysis, NZSSIGs are much underdeveloped. Very recently, a breakthrough has been achieved in [1], where the value functions and a Nash equilibrium of a very general class of NZSSIGs (assuming they exist and possess some regularity) have been characterized via a system of *quasi-variational inequalities* (QVIs). Besides the theoretical interest of this link by itself, it opens up the possibility of finding approximations to the value functions by numerically solving that system.

In this paper, we make the first attempt (to the best of our knowledge) at numerically solving NZSSIGs. In a nutshell, the iterative algorithm we put forward treats the NZSSIG at each iteration as a combination of a fixed point problem and a slowly relaxing one-player SIG. This allows us to take advantage of the machinery for the latter. Because, at the moment, we lack a proof of convergence, the proposed algorithm is admittedly heuristic. Instead, we report on a range of numerical experiments, which show convergence of the error with respect to the discretisation. In fact, errors are quite satisfactory and—in the examples we have tackled—the relative errors easily drop well below 0.1%. The algorithm can thus be used to assist further development of the field, as well as to gain insight into applications modelled by NZSSIGs. For even crafted NZSSIGs with exact solution are hard to construct: in fact, the authors of [1] also provide (as far as we know) the first example in the literature—thanks to which we were able to validate our algorithm in the first place. (For the sake of completeness, this solution has been included in Section 1.3.) Moreover, this example also evidences that

even when an analytical solution is available, its computation in practice may involve parameters which require solving complex nonlinear systems of equations—thus stressing the convenience of numerical approximations.

Let us outline the organization of the rest of the paper. We start in Section 2 by properly setting up NZSSIGs and recalling the main result of [1], namely, the Verification Theorem with the corresponding system of QVIs. For the sake of illustration, the concrete application of competition in energy markets is presented according to this framework. In Section 3, we briefly review the state-of-the-art numerics for one-player SIGs, with an emphasis on Howard's algorithm, which will later become a pivotal ingredient of our own numerical method for NZSSIGs. The latter is motivated, listed, and discussed in Section 4. Section 5 presents numerical evidence supporting it, in the light of which we draw some conclusions in Section 6.

1. The theoretical model

1.1. **Two-player nonzero-sum stochastic impulse games**

In this section we introduce a general class of two-player NZSSIG, to which our numerical algorithm is applicable. For the sake of briefness and pertinence, we will skip some of the most technical matters of the rigorous construction of the model. We refer the interested reader to [1] for more details and a more general modelling framework.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space under the usual conditions, that supports a standard one-dimensional Brownian motion $W = (W_t)$. For each $x \in \mathbb{R}$ we consider a process X (the *state variable*) starting at x,

$$X_t = x + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s + \sum_{k:\ \tau_k^1 \leq t} \delta_k^1 + \sum_{k:\ \tau_k^2 \leq t} \delta_k^2, \tag{1}$$

where:

(i) $\mu, \sigma : \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous.
(ii) $u^i := \{(\tau_k^i, \delta_k^i)\}_{k=1}^\infty$ $(i = 1, 2)$ is the *control* of player i, which consists of a sequence of stopping times τ_k^i (the k-th *intervention time* of player i) and $\mathcal{F}_{\tau_k^i}$-measurable random variables δ_k^i (the k-th *intervention impulse* of player i).

In words, when none of the players intervenes the state variable behaves as an Itô diffusion. The players can at any time decide to shift this process by applying a certain impulse. The specific type of interventions are determined by *(threshold-type) strategies*, which means that each player acts only when the state variable exits a given region of $\mathbb{R}$. More specifically, the control of player i is defined in terms of a strategy $\varphi_i := (\mathcal{C}_i, \xi_i)$, where $\mathcal{C}_i \subseteq \mathbb{R}$ is an open set (the *continuation region*) and $\xi_i : \mathcal{C}_i^c \to \mathbb{R}$ is a continuous function.[1] Player i intervenes if X exits $\mathcal{C}_i$—i.e., if for some $t \geq 0$ and $\omega \in \Omega$ it holds $X_t(w) \notin \mathcal{C}_i$—by applying an impulse $\xi_i(X_t(w))$. We decree that the game never stops and if both players want to intervene at the same time, then player 1 has the priority. The latter assumption is a matter of convention and not very restrictive.[2]

Henceforth, we write $\mathbb{E}_x$ to denote the conditional expectation given $X_{0-} = x$, and define the *objective function* for player i given the strategies (φ_1, φ_2) and the starting point x of X as

$$J^i(x; \varphi_1, \varphi_2) := \mathbb{E}_x \left[\int_0^\infty e^{-\rho_i s} f_i(X_s)ds + \sum_{k=1}^\infty e^{-\rho_i \tau_k^i} \phi_i\big(X_{(\tau_k^i)-}, \delta_k^i\big) + \sum_{k=1}^\infty e^{-\rho_i \tau_k^j} \psi_i\big(X_{(\tau_k^j)-}, \delta_k^j\big) \right], \tag{2}$$

with $i, j \in \{1, 2\}$, $j \neq i$; where:

(i) $\rho_i > 0$ is the subjective *discount rate* of player i.

---

[1]For any $A \subseteq \mathbb{R}$, A^c denotes the complement of A in $\mathbb{R}$.

[2]Under the assumptions of the Verification Theorem 1.2.1, the value of the game and the Nash equilibrium found would not change if player 2 had the priority instead. This can be checked by taking a pair $(\tilde{V}_1, \tilde{V}_2)$ as in the theorem and noticing that $(\tilde{V}_2, \tilde{V}_1)$ solves the game in which the structures (costs, gains, rates and payoffs) of the players have been swapped without swapping the priority.

(ii) $f_i : \mathbb{R} \to \mathbb{R}$ is a continuous function giving the *running payoff* of player i.

(iii) $\phi_i : \mathbb{R}^2 \to \mathbb{R}$ (resp. $\psi_i : \mathbb{R}^2 \to \mathbb{R}$) is a continuous function giving the *cost of intervention* (resp. *gain due to the opponent's intervention*) for player i.

In addition, we will only consider pairs of strategies (φ_1, φ_2) such that the previous expectations are well defined for all $x \in \mathbb{R}$, and we refer to these pairs as *admissible* (see [1] for more details).[3] Note for example that, for payoffs with polynomial growth, the 'no intervention strategies' $\phi_1 = \phi_2 = (\mathbb{R}, \emptyset)$ are always admissible. In this context, we say that the players behave optimally if their strategies form a Nash equilibrium. We recall that $(\varphi_1^*, \varphi_2^*)$ is a *Nash equilibrium* if it is an (admissible) pair of strategies such that, for all (φ_1, φ_2),

$$J^1(x; \varphi_1^*, \varphi_2^*) \geq J^1(x; \varphi_1, \varphi_2^*) \quad \text{and} \quad J^2(x; \varphi_1^*, \varphi_2^*) \geq J^2(x; \varphi_1^*, \varphi_2),$$

i.e., if player i changes strategy while player j does not, then on average player i will be worse off, and vice versa. If a Nash equilibrium $(\varphi_1^*, \varphi_2^*)$ exists, we define the *value function* for player i when using the strategies $(\varphi_1^*, \varphi_2^*)$ by

$$V_i(x) := J^i(x; \varphi_1^*, \varphi_2^*), \qquad i = 1, 2. \tag{3}$$

Our aim is to compute (V_1, V_2) for some Nash equilibrium, and more importantly, to retrieve from these values the equilibrium itself.

1.2. **The system of quasi-variational inequalities**

In order to establish a system of QVIs for (V_1, V_2) we need to define one last ingredient, known as the *intervention operators*, which will display the effect of each player's intervention on the value functions. For any two arbitrary functions $\tilde{V}_1, \tilde{V}_2 : \mathbb{R} \to \mathbb{R}$, $i, j \in \{1, 2\}$, $i \neq j$, and $x \in \mathbb{R}$, the *loss operator* of player i is given by

$$\mathcal{M}_i \tilde{V}_i(x) := \sup_{\delta \in \mathbb{R}} \{\tilde{V}_i(x + \delta) + \phi_i(x, \delta)\}. \tag{4}$$

If $\tilde{V}_i = V_i$ this operator gives the recomputed present value of i due to the cost of her own intervention. If for each $x \in \mathbb{R}$ there exists a unique $\delta^j(x) = \delta^j_{\tilde{V}_j}(x)$ that realizes the supremum in (4) , we also define the *gain operator* of player i as

$$\mathcal{H}_i \tilde{V}_i(x) := \tilde{V}_i(x + \delta^j(x)) + \psi_i(x, \delta^j(x)). \tag{5}$$

If $\tilde{V}_i = V_i$ this operator gives the recomputed present value of player i due to her opponent's intervention. Whenever we make use of a gain operator, we are implicitly stating that the above assumptions and notations are in place. We also emphasize that $\mathcal{H}_i \tilde{V}_i(x)$ depends on the whole function $\tilde{V}_j$ through $\delta^j_{\tilde{V}_j}$ and we will write $\mathcal{H}_i(\tilde{V}_j)\tilde{V}_i(x)$ instead, when we want to make this dependence explicit.

We can now state the Verification Theorem, due to [1], that will allow us to tackle the problem of finding (V_1, V_2) and a Nash equilibrium by numerically solving a deterministic system of QVIs. In the next theorem we use the notation $\mathcal{A}$ for the infinitesimal generator of X when no interventions take place. That is, $\mathcal{A}$ is the operator such that if $g \in C^2(S)$ for some $S \subseteq \mathbb{R}$, then

$$\mathcal{A}g(x) = \mu(x)g'(x) + \frac{1}{2}\sigma^2(x)g''(x), \quad \text{for } x \in S. \tag{6}$$

We concur that whenever we apply $\mathcal{A}$ to some function g, we are implicitly stating this function g is C^2 at every x at which we compute $\mathcal{A}g(x)$.

---

[3]In [1] admissibility is defined pointwise for greater generality. We refrain from doing this for simplicity.

Theorem 1.2.1 (System of QVIs). *Let* $\tilde{V}_1, \tilde{V}_2 : \mathbb{R} \to \mathbb{R}$ *such that for any* $i, j \in \{1, 2\}$, $i \neq j$:

$$\begin{cases} \mathcal{M}_j \tilde{V}_j - \tilde{V}_j \leq 0, & in \ \mathbb{R}, \\ \mathcal{H}_i \tilde{V}_i - \tilde{V}_i = 0, & in \ \{\mathcal{M}_j \tilde{V}_j - \tilde{V}_j = 0\}, \\ \max\left\{\mathcal{A}\tilde{V}_i - \rho_i \tilde{V}_i + f_i, \mathcal{M}_i \tilde{V}_i - \tilde{V}_i\right\} = 0, & in \ \{\mathcal{M}_j \tilde{V}_j - \tilde{V}_j < 0\} =: \mathcal{C}_j, \end{cases} \tag{7}$$

and $\tilde{V}_i \in C^2(\mathcal{C}_j \backslash \partial \mathcal{C}_i) \cap C^1(\mathcal{C}_j) \cap C(\mathbb{R})$ *has polynomial growth and bounded second derivative on some reduced neighbourhood of* $\partial \mathcal{C}_i$. *Suppose further that* $(\varphi_1^*, \varphi_2^*)$, *with* $\varphi_i^* := (\mathcal{C}_i, \delta_{\tilde{V}_i}^i)$, *is an admissible pair of strategies. Then*

$$(\tilde{V}_1, \tilde{V}_2) = (V_1, V_2) \ and \ (\varphi_1^*, \varphi_2^*) \ is \ a \ Nash \ equilibrium.$$

1.3. Example of application: competition in energy markets

We finish this Section by providing a specific example of application: competition in energy markets [1].

Let process X model the forward price of energy, evolving as a Brownian motion when there are no interventions. Consider the following two players: player 1 is an energy producer with unitary production cost s_1, so that in a simplified model her payoff is $X - s_1$. Player 2 runs a large company that buys from player 1 and sells at a unitary price $s_2 > s_1$, with payoff $s_2 - X$. Because of the high consumption of player 2, she can also affect the price of energy in the same way as player 1. Whenever these players intervene they incur a cost (advertising among other factors) which is modelled linearly for tractability—one constant component and another proportional to the change induced in the energy price. At the same time, because of their impact in the economy, the government subsidizes upon adverse movements in the energy price. For each player this represents a gain when the opponent intervenes, which is modelled in the same way as the cost of intervention, but with different parameters. We further assume that both players discount their winnings/losses at the same rate $\rho > 0$ and they have the same cost/gain parameters. More specifically,

$$X_t = x + \sigma W_t + \sum_{k: \ \tau_k^1 \leq t} \delta_k^1 + \sum_{k: \ \tau^2 \leq t} \delta_k^2$$

and for $i, j \in \{1, 2\}$, $i \neq j$,

$$J^i(x; \varphi_1, \varphi_2) := \mathbb{E}_x\left[\int_0^\infty e^{-\rho s}(-1)^{i-1}(X_s - s_i)ds - \sum_{k=1}^\infty e^{-\rho \tau_k^i}(c + \lambda|\delta_k^i|) + \sum_{k=1}^\infty e^{-\rho \tau_k^j}(\tilde{c} + \tilde{\lambda}|\delta_k^j|)\right], \tag{8}$$

where $0 \leq \tilde{c} \leq c$, $0 \leq \tilde{\lambda} \leq \lambda$, $(c, \lambda) \neq (\tilde{c}, \tilde{\lambda})$, and $1 - \rho\lambda > 0$. These parametric restrictions ensure among other things the existence of a Nash equilibrium (see [1] for more details).

The exact solution to this game can be found by application of Theorem 1.2.1. To this purpose, the system of QVIs (7) is heuristically solved, yielding candidates for value functions and a Nash equilibrium. This is done, first, by making some educated guesses regarding the shape of the optimal continuation regions and value functions. Second, by solving the ordinary differential equations (ODEs) in (7) where appropriate. Finally, by imposing the regularity requirements of Theorem 1.2.1 through pasting conditions. Upon verification of the remaining hypotheses, the following turns out to be the solution to the game:

$$V_2(x) = \begin{cases} \varphi^{A_1, A_2}(x_1^*) + \tilde{c} + \tilde{\lambda}(x_1^* - x) & \text{if } x \in (-\infty, \bar{x}_1], \\ \varphi^{A_1, A_2}(x) & \text{if } x \in (\bar{x}_1, \bar{x}_2), \qquad V_1(x) = V_2(2\tilde{s} - x), \\ \varphi^{A_1, A_2}(x_2^*) - c - \lambda(x - x_2^*) & \text{if } x \in [\bar{x}_2, +\infty), \end{cases}$$

where:

$$\varphi^{A_1, A_2}(x) = A_1 e^{\theta x} + A_2 e^{-\theta x} + \frac{1}{\rho}(s_2 - x),$$

$$\tilde{s} := \frac{s_1 + s_2}{2}, \qquad \theta := \sqrt{2\rho/\sigma^2}, \qquad \eta := (1 - \lambda\rho)/\rho,$$

$$\bar{x}_i := \tilde{s} + \frac{(-1)^i}{\theta} \log\left(\sqrt{\frac{\eta + \xi}{\eta - \xi}}\left(\sqrt{\Gamma + 1} + \sqrt{\Gamma}\right)\right), \quad x_i^* := \tilde{s} + \frac{(-1)^i}{\theta} \log\left(\sqrt{\frac{\eta - \xi}{\eta + \xi}}\left(\sqrt{\Gamma + 1} + \sqrt{\Gamma}\right)\right),$$

$$A_i := \exp\left((-1)^i \theta \tilde{s}\right) \frac{\sqrt{\eta^2 - \xi^2}}{2\theta}\left((-1)^{i+1}\sqrt{\Gamma + 1} - \sqrt{\Gamma}\right),$$

$$\Gamma := \frac{\theta(c - \tilde{c})}{4\xi} + \frac{\theta c(\lambda - \tilde{\lambda})}{4\eta\xi} + \frac{\lambda - \tilde{\lambda}}{2\eta}$$

and $\xi \in (0, \eta)$ is the unique zero of $F(y) := 2y - \eta \log\left(\frac{\eta + y}{\eta - y}\right) + \theta c$.

As it can be readily noticed, the analytical solution involves the computation of several parameters and the resolution of at least one nonlinear equation. As a matter of fact, the number of parameters in this solution was reduced making use of the symmetry in the problem. In a more general two-player NZSSIG under the modelling framework of Section 1.1, if we assume the optimal continuation regions in Theorem 1.2.1 are semi-bounded intervals, then an analytical solution can easily involve eight parameters: two finite limits of both continuation regions, two maximizing points of the net value functions (subtracting intervention costs) and four undetermined constants coming from the second order ODEs. Moreover, these parameters are the solution to a nonlinear system of equations that arises from smooth pasting and optimality conditions [1, Def. 4.1.]. In the more frequent than not situation in which this system cannot be simplified, computing the analytical solution may become prohibitive. This further motivates the need for a numerical algorithm to solve the system of QVIs (7).

Remark. To the best of our knowledge this is, at the time of writing, the only one example in the literature of an analytically solvable NZSSIG. Henceforth, we will refer to it as the *benchmark game*.

2. Numerics for one-player SIGs: state of the art

Notation. Henceforward, finite grids will be denoted by $\mathbb{K}$ and $\mathbb{S}$. For a fixed grid, we will use the same font type when discretising an operator, e.g., $\mathcal{M}$, $\mathcal{H}$ and $\mathcal{A}$ will be replaced by $\mathbb{M}$, $\mathbb{H}$ and $\mathbb{A}$. We shall specify later the way the discretisations have been done in our experiments. We will not change notation for functions over grids as their domain will always be clearly stated. The only exception will be those functions that have been redefined to account for boundary conditions (BCs). Lastly, we recall that for any set A, $\mathbb{R}^A$ denotes the set of functions from A to $\mathbb{R}$.

Policy iteration or *Howard's* algorithm for (discrete) variational inequalities (VIs) was originally developed in [3, 4, 9]. The method was then extended to QVIs [7, 8, 10], i.e., variational inequalities in which the obstacle depends on the solution itself. Let $\mathbb{K} \subset \mathbb{R}$ be a finite set (the grid). These problems have the form:

$$\text{Find } V \in \mathbb{R}^{\mathbb{K}}: \quad \max\left\{\mathbb{L}V(x) + \mathbb{g}(x), \max_{y \in \mathbb{K}} \mathbb{B}^y V(x) - V(x)\right\} = 0 \quad \text{for all } x \in \mathbb{K},^{[4]} \tag{9}$$

where $\mathbb{g} \in \mathbb{R}^{\mathbb{K}}$ and $\mathbb{L}, \mathbb{B}^y : \mathbb{R}^{\mathbb{K}} \to \mathbb{R}^{\mathbb{K}}$ are linear and affine operators resp., for each $y \in \mathbb{K}$.

Problem 9 encomprises in particular a discrete localized version of a one-player SIG (i.e., a stochastic impulse control problem) since the system (7) reduces in this case to one single QVI. Indeed, by an appropriate finite difference approximation (more details in Section 3) one can take $\mathbb{g} = f|_{\mathbb{K}}$ (modified to account for Dirichlet or Neumann BCs),[5] $\mathcal{A} - \rho_1 Id \approx \mathbb{L}$ and $\mathcal{M}_1 \tilde{V}(x) \approx \max_{y \in \mathbb{K}} \mathbb{B}^y \tilde{V}(x)$, with $\mathbb{B}^y \tilde{V}(x) := \tilde{V}(y) + \phi_1(x, y - x)$. By

---

[4]Slightly more general formulations are also available.

[5]For any function $F : A \to B$ and $X \subseteq A$, $F|_X$ denotes the restriction of F to X.

their antisymmetric structure, two-player zero-sum SIGs can also be accommodated with a similar formulation as a max-min double obstacle problem [1, page 7] for one single value function $V := V_1 = -V_2$, and the policy iteration algorithm can be adapted [5].

However, this framework is not general enough to tackle the NZSSIGs described in Section 1.1 and the full system of QVIs (7). There are, to the best of our knowledge, no available numerical methods to approach the latter problem. We present now the classical policy iteration algorithm used to solve (9). It will be embedded in the numerical scheme we will put forward to solve the much more general problem (7). While we note that our later use of this algorithm is not fully within the theoretical scope studied in [7]—in particular, we will need to relinquish the affine nature of the operators $\mathbb{B}^y$, $y \in \mathbb{K}$—the two following main assumptions will be in place nonetheless ($\#\mathbb{K}$ is the cardinal of set $\mathbb{K}$):

Assumptions 2.1.1. *(i) $-\mathbb{L}$ is a strictly diagonally dominant M-operator. That is, if $L \in \mathbb{R}^{\#\mathbb{K} \times \#\mathbb{K}}$ is the canonical matrix of $\mathbb{L}$,[6] then*

$$L_{ij} \geq 0 \text{ for all } i \neq j \text{ and } -L_{ii} > \sum_{j \neq i} L_{ij} \text{ for all } i.$$

(ii) $\mathbb{B}^y$ is a non-expansive function for $\|\cdot\|_\infty$, for all $y \in \mathbb{K}$. That is,

$$\|\mathbb{B}^y \tilde{V}_1 - \mathbb{B}^y \tilde{V}_2\|_\infty \leq \|\tilde{V}_1 - \tilde{V}_2\|_\infty, \quad \text{for all } \tilde{V}_1, \tilde{V}_2 \in \mathbb{R}^\mathbb{K}, \ y \in \mathbb{K}.$$

We remark that, in particular, Assumption *(i)* will guarantee that Algorithm 1 is well defined, as the linear operators $\mathbb{L}^k$ will be non-singular [5].

Algorithm 1 Policy iteration for one QVI (one-player SIG)

1: Set $\varepsilon > 0$ (numerical tolerance) and $k_{\max} \in \mathbb{N}$ (maximum iterations).
2: Pick initial guess: $V^0 \in \mathbb{R}^\mathbb{K}$.
3: Let $k = 0$ (iteration counter) and $R^0 = +\infty$.
4: **while** $R^k > \varepsilon$ **and** $k \leq k_{\max}$ **do**
5:     $M^k := \max_{y \in \mathbb{K}} \mathbb{B}^y V^k$.
6:     $\alpha^k := \mathbb{1}_{\{\mathbb{L} V^k + \mathrm{g} < M^k - V^k\}}$ (action at each point).
7:     Define $\mathbb{L}^k : \mathbb{R}^\mathbb{K} \to \mathbb{R}^\mathbb{K}$ and $\mathrm{g}^k \in \mathbb{R}^\mathbb{K}$ by

$$\mathbb{L}^k \tilde{V}(x) := \begin{cases} \mathbb{L}\tilde{V}(x) & \text{if } \alpha^k(x) = 0 \\ -\tilde{V}(x) & \text{if } \alpha^k(x) = 1 \end{cases} \qquad \mathrm{g}^k(x) := \begin{cases} \mathrm{g}(x) & \text{if } \alpha^k(x) = 0 \\ M^k(x) & \text{if } \alpha^k(x) = 1. \end{cases}$$

8:     Solve for V^{k+1}: $\mathbb{L}^k V^{k+1} + \mathrm{g}^k = 0$.
9:     $R^{k+1} := \|V^{k+1} - V^k\|$.
10:    $k = k + 1$.
11: **end while**

Note that in Algorithm 1, computing the operators $\mathbb{L}^k$ amounts simply to redefine a matrix row by row using either the matrix of $\mathbb{L}$ or $-Id$. We have chosen an 'operators-type' notation for this paper, as it will simplify matters in the sequel, when compared with its matrix counterpart.

3. PROPOSED ALGORITHM FOR TWO-PLAYER NZSSIGS

Compared with the single-value-function problems in the previous Section, general two-player NZSSIGs are distinctly more challenging. The main challenges are:

---

[6]If $\mathbb{K}$ is the grid: $x_0 < x_1 < \cdots < x_M$, then L is the matrix with columns $L_j = \mathbb{L}\mathbb{1}_{\{x_j\}}$, for $j = 0, \ldots, M$.

- two value functions, governed by a system of QVIs, must be solved for,
- the dependence between (V_1, V_2) is highly nonlinear due to the presence of the gain operators $\mathcal{H}_i(V_j)V_i$,
- each gain operator is expansive as a function of $(\tilde{V}_1, \tilde{V}_2)$; and
- solutions will typically be less regular. For example, if (V_1, V_2) is the solution to the benchmark game in section 1.3 then V_i is singular at $\bar{x}_j$ for $i, j \in \{1, 2\}$, $i \neq j$ (i.e., each value function is non differentiable at the border of the opponent's continuation region), in spite of the game having linear payoffs, costs and gains. This is somehow to be expected, as Theorem 1.2.1 contemplates this lack of regularity within the smoothness assumptions (compare to the classical Verification Theorems for one-player problems [10] where greater regularity is assumed).

Algorithm 2 below is, as far as we know, the first ever numerical attempt at two-player NZSSIGs. At present, it is admittedly heuristic and supported only by the numerical evidence reported in Section 4.

The remainder of this Section is organized as follows. We start by explaining the idea and motivating the underlying heuristics. Then, we describe in detail the discretisation of the system of QVIs (7). Finally, we list the new algorithm.

Heuristics. Having discretised the system of QVIs (7), we start from an initial guess (V_1^0, V_2^0) to approximate its solution. We seek an iterative procedure to consecutively compute (V_1^{k+1}, V_2^{k+1}), given (V_1^k, V_2^k) at the k-th iteration.

Let $i, j \in \{1, 2\}$, $i \neq j$. A natural idea is, first, to partition the grid into the 'approximate continuation region' of player j—$\{\mathbb{M}_j V_j^k - V_j^k < 0\}$—and its complement, the 'approximate *intervention region*'; and then to compute V_i^{k+1} either by calculating a gain as $\mathbb{H}_i(V_j^k)V_i^k$ in the former case, or by solving one QVI with Howard's algorithm (Algorithm 1) in the latter. Note, however, that naively defining $\{\mathbb{M}_j V_j^k - V_j^k < 0\}$ as the 'approximate continuation region' of player j poses difficulties.

Indeed, the discretisation of the loss operator $\mathcal{M}_j$ as $\mathbb{M}_j$ implies that even the true value function V_j will generally not verify $\mathbb{M}_j V_j(x) - V_j(x) = 0$ (or $\mathbb{M}_j V_j(x) - V_j(x) \geq 0$) for x in the true intervention region of player j. We therefore need to relax this constraint to account both for numerical error and the discrepancy between the discrete and space-continuous problems. Our experiments have shown that a successive relaxation procedure turns out to be the most effective. Consequently, we define the approximate continuation region of player j as $\{\mathbb{M}_j V_j^k - V_j^k < -r^k\}$ instead, where r^k is a small positive number. By letting r^k relax to a preset small tolerance $\varepsilon > 0$, the iterative approximations will hopefully converge to the correct discrete solution.

It remains to schedule the relaxation procedure and to define a measure of convergence for the algorithm. Regarding the former, having computed (V_1^{k+1}, V_2^{k+1}), r^k is linearly relaxed (line 11). (This relaxation procedure is chosen for simplicity.). Then the largest pointwise residual to the system of QVIs (incurred by either approximate value function) is calculated across the grid (line 12), taking into account the numerical tolerance ε. We denote this residual R^{k+1} and we consider the algorithm has converged when it drops below a certain tolerance—unlike in Algorithm 1, where the residual is taken as the distance between consecutive approximations. This alternative approach has been chosen for being more informative. It reflects whether a solution to the discrete system of QVIs has been found—as opposed to the algorithm stagnating—on top of giving valuable information at each grid node.

By construction, the junctions between the approximate continuation and intervention regions for each of the players will necessarily take place on top of a finite difference node. This may lead to numerical issues when the exact value functions are non differentiable there (as will often be the case). Namely, the pointwise residual to the QVIs at the junction nodes may not be made arbitrarily small by refining the grid—because the derivatives contained in the residual are not defined at the exact junction, in the first place. In those cases, the algorithm must be stopped at that point, since adding more grid nodes in a naive way cannot lead to any improvement.

As a final remark, we mention that numerical experiments show Algorithm 2 does not enjoy global convergence (i.e. it is not guaranteed to converge from arbitrary initial guesses). Providing a good enough pair (V_1^0, V_2^0) is thus a practical issue; in Section 4, a natural way of constructing educated guesses is explained.

Discretisation. Let $\mathbb{S} \subseteq \mathbb{R}$ be a finite set. $\mathbb{S}$ is the grid we will use to discretise the system (7). In all of the numerical experiments described in the sequel we have taken $\mathbb{S}$ as an equispaced grid of M steps between certain $x_{\min} < 0 < x_{\max}$ with $|x_{\min}|, |x_{\max}|$ and M big enough (more about this below).

Let $i, j \in \{1, 2\}$, $i \neq j$, $\tilde{V}_1, \tilde{V}_2 \in \mathbb{R}^{\mathbb{S}}$ and $x \in \mathbb{S}$. We proceed to define the discretised versions of the loss and gain operators, over the grid $\mathbb{S}$, as

$$\mathbb{M}_i \tilde{V}_i(x) := \max_{y \in \mathbb{S}}\{\tilde{V}_i(y) + \phi_i(x, y - x)\} \qquad \mathbb{H}_i(\tilde{V}_j)\tilde{V}_i(x) := \tilde{V}_i(\mathrm{y}_j(x)) + \psi_i(x, \mathrm{y}_j(x) - x),$$

where

$$\mathrm{y}_j(x) := \min\left(\operatorname*{argmax}_{y \in \mathbb{S}}\{\tilde{V}_j(y) + \phi_j(x, y - x)\}\right).$$

Next, we choose a finite difference scheme for the ODEs

$$0 = \mathcal{A}\tilde{V}_i - \rho_i\tilde{V}_i + f_i = \frac{1}{2}\sigma^2\tilde{V}_i'' + \mu\tilde{V}_i' - \rho_i\tilde{V}_i + f_i$$

which is consistent, monotone and stable, adding Dirichlet or Neumann-type BCs. In all of our experiments, we have chosen an upwind finite difference scheme, where

$$\tilde{V}_i'(x) \approx \frac{\tilde{V}_i\big(x + \operatorname{sgn}(\mu(x))h\big) - \tilde{V}_i(x)}{\operatorname{sgn}(\mu(x))h} \qquad \tilde{V}_i''(x) \approx \frac{\tilde{V}_i(x + h) - 2\tilde{V}_i(x) + \tilde{V}_i(x - h)}{h^2}, [7]$$

and $\tilde{V}_i(x_{\min} - h), \tilde{V}_i(x_{\max} + h)$ were solved for using Neumann conditions on

$$\tilde{V}_i'\left(x_{\min} - \frac{1 + \operatorname{sgn}(\mu(x_{\min}))}{2}h\right), \quad \tilde{V}_i'\left(x_{\max} + \frac{1 - \operatorname{sgn}(\mu(x_{\max}))}{2}h\right)$$

respectively. How to get these conditions, and in particular how to choose $x_{\min}$ and $x_{\max}$, is a problem-specific question. In some situations, and particularly in the models numerically tested in this paper, one can heuristically assert that at a Nash equilibrium the continuation region of player i should be a semi-interval of the form $\mathcal{C}_i = (\underline{x}_i, \overline{x}_i)$—with one end-point finite and the other infinite. Intuitively, one can further guess that there should exist a unique $y_i^* \in (\underline{x}_i, \overline{x}_i)$ that maximizes the net value of player i when she intervenes (see, e.g., [1, Sect. 4.2] and [2, Sect. 2.3.1]). These conjectures arise mainly from the observation of the payoff functions f_1, f_2 which encode the goals of the players and roughly hint at some broad regions where each player would like the state variable to remain at. On $\mathbb{R}\backslash\mathcal{C}_i$ and $\mathbb{R}\backslash\mathcal{C}_j$ one of the two players will intervene, thus giving either

$$V_i(x) = \mathcal{M}_i V_i(x) = V_i(y_i^*) + \phi_i(x, y_i^* - x) \quad \text{or} \quad V_i(x) = \mathcal{H}_i V_i(x) = V_i(y_j^*) + \psi_i(x, y_j^* - x).$$

If on the interiors $(\mathbb{R}\backslash\mathcal{C}_i)^o$ and $(\mathbb{R}\backslash\mathcal{C}_j)^o$ the derivatives $\frac{d\phi_i(x, y_i^* - x)}{dx}$ and $\frac{d\psi_i(x, y_j^* - x)}{dx}$ exist and do not depend on y_i^*, y_j^* resp., differentiating the previous relations yields the Neumann conditions (provided $x_{\min}, x_{\max}$ are 'extreme' enough). The previous requirements on the derivatives are satisfied, for example, if the cost and gain structures have the form $\phi_i(x, \delta) = g_i(x) + a_i|\delta|$ and $\psi_i(x, \delta) = h_i(x) + b_i|\delta|$, for some differentiable functions $g_i, h_i : \mathbb{R} \to \mathbb{R}$ and $a_i, b_i \in \mathbb{R}$. For the benchmark game the Neumann BCs read $V_1'(x_{\min} - h) = \lambda$, $V_1'(x_{\max}) = \tilde{\lambda}$, $V_2'(x_{\min} - h) = -\tilde{\lambda}$ and $V_2'(x_{\max}) = -\lambda$. We will denote by $\mathbb{A}$ the discretised version of $\mathcal{A}$ and $\mathbb{f}_i$ the restriction of f_i to $\mathbb{S}$, redefined at $\min \mathbb{S}$ and $\max \mathbb{S}$ to account for the BCs.

For Algorithm 2 we recall that given any real number a, a^+ denotes its positive part (i.e., $a^+ := \max\{a, 0\}$) and for any subset $S \subseteq \mathbb{R}$, $\mathbb{1}_S$ denotes the indicator function of S.

---

[7]sgn denotes the sign function, i.e., $\operatorname{sgn}(x) = 1$ if $x \geq 0$ and $\operatorname{sgn}(x) = -1$ otherwise.

Algorithm 2 Policy iteration for system of QVIs (two-player NZSSIG)

1: Set $\varepsilon > 0$ (numerical tolerance), $0 < \alpha < 1$, $r^0 > 0$ (relaxation parameters) and $k_{\max} \in \mathbb{N}$ (maximum iterations).
2: Pick initial guess: $(V_1^0, V_2^0) \in \mathbb{R}^\mathbb{S} \times \mathbb{R}^\mathbb{S}$.
3: Let $k = 0$ (iteration counter) and $R^0 = +\infty$
4: **while** $R^k > \varepsilon$ **and** $k \leq k_{\max}$ **do**
5:     **for** i=1, 2 (player i) **do**
6:         $j = 3 - i$ (player j).
7:         $\mathcal{C}_j^k := \{\mathbb{M}_j V_j^k - V_j^k < -r^k\}$.
8:         For $x \notin \mathcal{C}_j^k$, let $V_i^{k+1}(x) = \mathbb{H}_i(V_j^k)V_i^k(x)$.
9:         For $x \in \mathcal{C}_j^k$, solve for $V_i^{k+1}(x)$ by applying Algorithm 1 to
           $$\max\left\{\mathbb{A}V_i^{k+1}(x) - \rho_i V_i^{k+1}(x) + \mathbb{f}_i(x),\ \mathbb{M}_i V_i^{k+1}(x) - V_i^{k+1}(x)\right\} = 0.$$
10:     **end for**
11:     $r^{k+1} := \max\{\alpha r^k, \varepsilon\}$ (relaxation).
12:     Let R^{k+1} be the largest pointwise residual to the system of QVIs, i.e.

$$R^{k+1} := \max_{\substack{i,j\in\{1,2\}, j\neq i \\ x\in\mathbb{S}}}\left\{\left(\mathbb{M}_i V_i^{k+1}(x) - V_i^{k+1}(x)\right)^+, \right. \tag{10}$$

$$\left|\mathbb{H}_i(V_j^{k+1})V_i^{k+1}(x) - V_i^{k+1}(x)\right|\mathbb{1}_{\mathcal{C}_j^{k,\varepsilon}}(x),$$

$$\left.\left|\max\left\{\mathbb{A}V_i^{k+1}(x) - \rho_i V_i^{k+1}(x) + \mathbb{f}_i(x), \mathbb{M}_i V_i^{k+1}(x) - V_i^{k+1}(x)\right\}\right|\mathbb{1}_{\mathbb{S}\setminus\mathcal{C}_j^{k,\varepsilon}}(x)\right\},$$

        where $\mathcal{C}_j^{k,\varepsilon} := \{\mathbb{M}_j V_j^{k+1} - V_j^{k+1} < -\varepsilon\}$.
13:     Let $k = k + 1$.
14: **end while**

Line 9 of Algorithm 2 deserves some special attention. Although at this step we want to solve a problem restricted to the subgrid $\mathbb{K} = \mathcal{C}_j^k$ (for fixed k, i, j), we still need the information in $\mathbb{S}\setminus\mathcal{C}_j^k$ in two ways, namely:

- To compute the non-local operators $\mathbb{M}_i$.
- To restrict to $\mathbb{K}$ the equation $\mathbb{A}V_i^{k+1}(x) - \rho_i V_i^{k+1}(x) + \mathbb{f}_i(x) = 0$, properly accounting for the original BCs.

Suppose $\mathbb{S}$ is the grid $x_0 < x_1 < \cdots < x_M$ and let us identify each point with its respective index. Let $A \in \mathbb{R}^{(M+1)\times(M+1)}$ be the canonical matrix of $\mathbb{A}$. For any subsets $I, J \subseteq \{0, \ldots, M\}$, let us write $A_{I,J}$ for the submatrix of A which has rows indexed in I and columns indexed in J, and $\mathbb{A}_{I,J}$ the associated operator. Put $H_i^k := V_i^{k+1}|_{\mathbb{S}\setminus\mathbb{K}} = \mathbb{H}_i(V_j^k)V_i^k|_{\mathbb{S}\setminus\mathbb{K}}$ and $h_i^k := \max H_i^k$. Then in order to apply Algorithm 1 we take

$$\mathbb{L}\tilde{V} = \mathbb{A}_{\mathbb{K},\mathbb{K}}\tilde{V} - \rho_i\tilde{V}, \quad \mathbb{g} = \mathbb{f}_i|_\mathbb{K} + \mathbb{A}_{\mathbb{K},\mathbb{S}\setminus\mathbb{K}}H_i^k \quad \text{and} \quad \mathbb{B}^y\tilde{V}(x) = \max\{\tilde{V}_i(y) + \phi_i(x, y - x), h_i^k\},$$

for all $\tilde{V} \in \mathbb{R}^\mathbb{K}$, $x, y \in \mathbb{K}$. We remark once again that the functions $\mathbb{B}^y$ fail to be affine operators. This, together with some assumptions which are not satisfied, make this application of Algorithm 1 fall outside of the scope of [7] (and even more of [8, 10]). However, it is easy to check that the main assumptions, Assumptions 2.1.1, are still verified under our discretisation and we have observed unconditional convergence of this subroutine (to very high orders of precision) in all the experiments we have ran.

Lastly, we note that Algorithm 1 requires setting a numerical tolerance which needs not be the same as the one of Algorithm 2. In fact, we have always taken it strictly smaller given the perceived unconditional convergence of the former and its precision.[8]

4. Numerical results

In this section we assess the performance of Algorithm 2. We shall call the problems considered in the experiments simply 'games'. The two value functions V_1, V_2 are approximated as described in Section 3 on an equispaced grid of $M + 1$ nodes, and the computational domain is the plotted one. In all the subsequent games, $\varepsilon = 10^{-8}$, $\alpha = 0.8$ and $r_0 = 1$. The largest pointwise residual (equation (10)) at convergence is denoted by R^∞.

As mentioned before, the only NZSSIG for which an analytical solution is currently available is the benchmark game in Section 1.3. Therefore, we shall eventually focus on that problem. However, we introduce two other games first (for which we do not have an analytical solution). They illustrate how the initial guess for the benchmark game can be constructed and provide further numerical evidence supporting Algorithm 2.

4.1. Parabolic game

This is a version of the benchmark game where the payoff is replaced by a concave parabola with roots r_i^L and r_i^R:

$$\hat{f}_i(x) := -(x - r_i^L)(x - r_i^R), \qquad r_i^L < r_i^R. \tag{11}$$

Let us motivate this game. We seek to use as initial guess the value functions of the 'unilateral games', that is, the control problems in which one of the players never intervenes (her continuation region is fixed and equal to $\mathbb{R}$). Removing the action of one of the players, however, may not always lead to a well-posed problem. Indeed, for payoffs without maximum one could end up for example with 'infinite-valued value functions' or 'infinite-valued optimal impulses'. This is indeed the case for the benchmark game. Thus, in order to skirt that difficulty it is convenient to define variations of the benchmark with payoffs that attain a maximum, like (11). Note that, when well-posed, the unilateral games can be readily solved with Howard's algorithm (Algorithm 1).

Figure 1 shows the numerical solution, (V_1, V_2), to a parabolic game (pair of solid curves) along with the initial guess (pair of dashed curves). The latter are, in turn, the value functions of the corresponding unilateral games for players 1 and 2. It is intuitively clear that (V_1, V_2) approximate, over the grid, functions which indeed satisfy the assumptions of the Verification Theorem 1.2.1. The Nash equilibrium exhibited in this Theorem— $(\varphi_1^*, \varphi_2^*)$—can be retrieved from this graph. Indeed, note that in this case the cost for player i is $\phi_i = -100$ and her approximate continuation region is $\{\mathbb{M}_i V_i - V_i < \varepsilon\} = \{\max V_i - V_i - 100 < \varepsilon\}$. Further, her optimal impulse at a given x is the one that realizes (the discrete analogue of) the supremum in equation (4), which in this case amounts to translating x to the maximizing point of V_i. Consequently, we get $\varphi_1^* = \big((-\infty, 1.068), -1.848 - x\big)$ and $\varphi_2^* = \big((-3.048, +\infty), -0.120 - x\big)$.

---

[8]Note that although precision is assessed in a different way in Algorithm 1, the pointwise residuals to the QVIs on the corresponding regions are afterwards checked as part of Algorithm 2
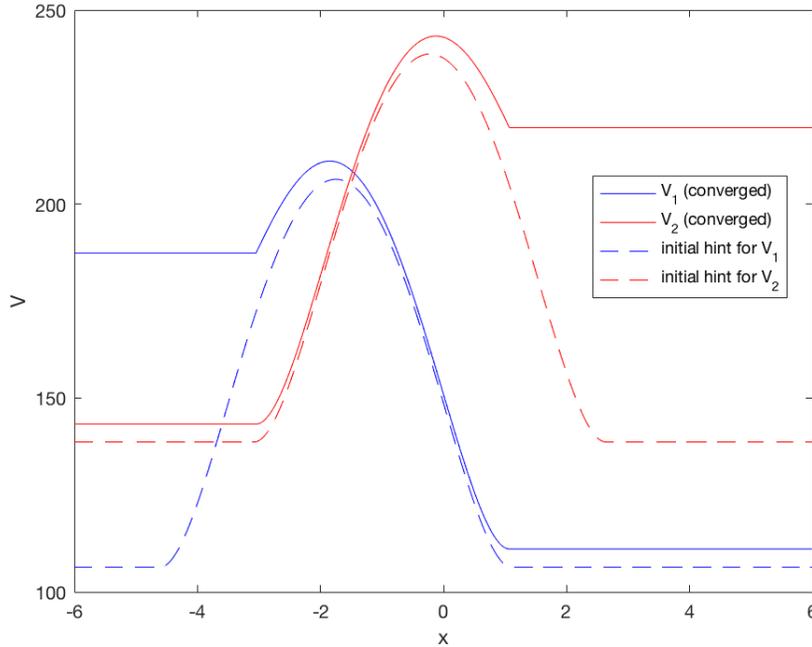
FIGURE 1. Value functions and Nash equilibrium with parabolic running payoffs: $\hat{f}_1 = -(x + 4.5)(x - 1)$ and $\hat{f}_2 = -(x + \pi)(x - 2.7)$. Overlaid, initial guesses (solutions of the respective unilateral games) Parameters: $\rho = 0.03$, $\sigma = 0.25$, $c = 100$, $\tilde{c} = 30$, $\lambda = \tilde{\lambda} = 0$. Here, $M = 1000$; check Table 1 for other values.

As it turns out, this parabolic game also converges from the 'zero guess' ($V_1^0 = V_2^0 = 0$). On Table 1, the convergence of R^∞ over a wide range of finite difference grids (with $301, 601, \ldots, 3001$ nodes) is compiled. In all cases, both initial guesses lead to convergence (to within ε) of Algorithm 2. Nonetheless, the algorithm takes in general fewer iterations when starting from the values of the unilateral parabolic games. (We stress that those on Table 1 are the outer iterations of Algorithm 2. Within each one there is an inner loop of Howard's iterations. Thus the computational cost scales linearly with the number of outer iterations.)

M	R^∞	its.[a]	its.[b]
300	1.4×10^{-12}	54	53
600	4.0×10^{-9}	74	77
1200	3.3×10^{-9}	144	77
1800	9.7×10^{-9}	95	77
2400	7.3×10^{-9}	123	77
3000	5.9×10^{-9}	215	103

TABLE 1. Parabolic game: largest residual to QVIs at convergence (R^∞) vs. grid points ($M + 1$). Iterations to convergence within $\varepsilon = 10^{-8}$ starting from: zero guess (its.[a]) and value functions for one-player games (its.[b]). (Same parameters as in Figure 1.)

Since the exact solution to the parabolic game is unknown, we cannot say anything about the convergence of Algorithm 2 in continuous-space. On the other hand, we see the system of QVIs is (approximately) enforced

to within $\varepsilon \ll 1$ by a pair of numerical functions which—also approximately—comply with the regularity assumptions of Theorem 1.2.1. As such, they are the (approximate) value functions of the parabolic game.

We conclude this example by illustrating the interplay between the value functions which numerically solve the system of QVIs, the Nash equilibrium derived from them, and the evolution of the optimally controlled underlying. Once the optimal strategies $(\varphi_1^*, \varphi_2^*)$ are available, they can be executed on specific realizations of the game. Sticking to the parameters and numerical solution in Figure 1, Figure 2 depicts one exemplary trajectory of the underlying in the time interval $0 \le t \le 1000$, starting from $x = 0$ and subjected to the pair of optimal strategies. For numerical purposes, let us define

$$\hat{J}_T^i(x; \varphi_1, \varphi_2) := \mathbb{E}_x \left[\int_0^T e^{-\rho_i s} f_i(X_s) ds + \sum_{k:\ \tau_k^i \le T} e^{-\rho_i \tau_k^i} \phi_i\left(X_{(\tau_k^i)^-}, \delta_k^i\right) + \sum_{k:\ \tau_k^j \le T} e^{-\rho_i \tau_k^j} \psi_i\left(X_{(\tau_k^j)^-}, \delta_k^j\right) \right]. \quad (12)$$
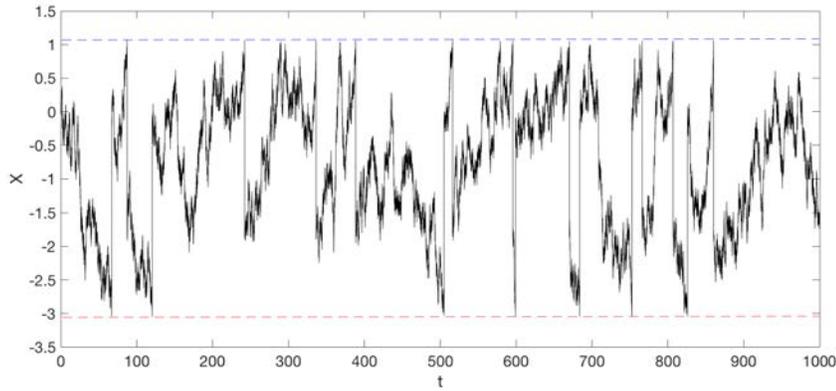

FIGURE 2. Exemplary trajectory from $x = 0$ (parameters and solution from Figure 1). The blue and red dashed lines are the intervention thresholds for players 1 and 2, respectively.
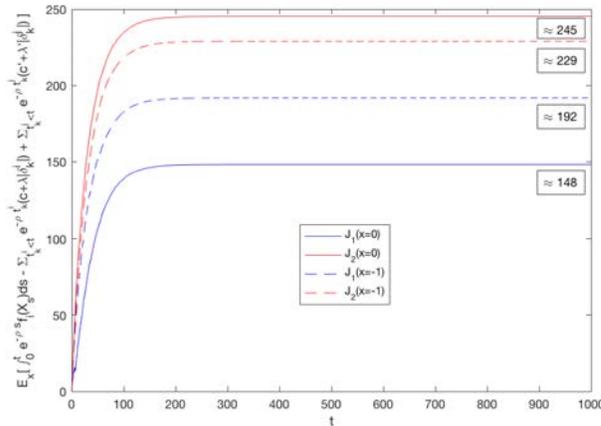

FIGURE 3. Approximations to the objective functions at $x = 0$ (solid curves) and $x = -1$ (dashed curves), obtained by Monte Carlo simulation (see text for details). Parameters and optimal strategies from Figure 1. Compare with $V_1(0)$, $V_2(0)$, $V_1(-1)$ and $V_2(-1)$ there.

Intuitively, $\hat{J}_T^i(x; \varphi_1, \varphi_2) \to J^i(x; \varphi_1, \varphi_2)$ (defined in (2)) as $T \to \infty$. In fact, after $T \gtrsim 300$, the integrals in (12) for the parabolic game have essentially attained their asymptotic value. Thus, we simply take $J^i(x; \varphi_1, \varphi_2) \approx \hat{J}_{T=1000}^i(x; \varphi_1, \varphi_2)$. With this clarification, Figure 3 shows in particular $V_1(0) = J^1(0; \varphi_1^*, \varphi_2^*)$, $V_2(0) = J^2(0; \varphi_1^*, \varphi_2^*)$, $V_1(-1) = J^1(-1; \varphi_1^*, \varphi_2^*)$, and $V_2(-1) = J^2(-1; \varphi_1^*, \varphi_2^*)$, obtained by Monte Carlo simulation.[9] They compare fairly well with the values of $V_1(0)$, $V_2(0)$, $V_1(-1)$ and $V_2(-1)$ in Figure 1 obtained by our algorithm. (Even better agreement could be obtained by increasing M in that figure and reducing the discretisation bias and statistical error of the Monte Carlo simulation, but this is good enough to make our point.)
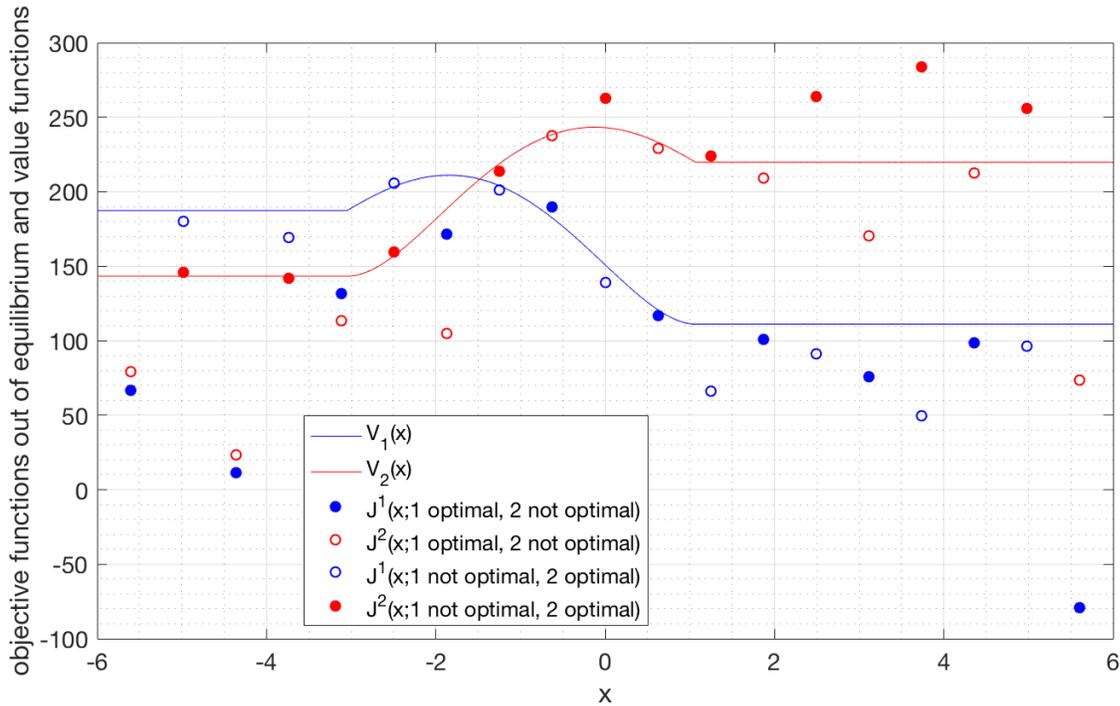

FIGURE 4. Empty blue [red] circles mean that the player 1 [2] has departed from her optimal strategy while her opponent has not. By virtue of Nash equilibrium, she cannot be (within numerical errors) better off than $V_1(x)$ [$V_2(x)$]. Full circles: objective function of the player who does not drop her optimal strategy. Parameters and optimal strategies from Figure 1. (Note that results are subjected to numerical error.)

The Nash equilibrium itself can be visually explored in the following way. For a given starting point x, we keep the optimal strategy for one of the two players, and slightly alter the strategy of the other one. For concreteness, let us assume that player 1 "moves" (i.e. $\varphi_1 = (1 \pm 0.25\mathcal{U})\varphi_1^*$) while player 2 does not ($\varphi_2 = \varphi_2^*$). ($\mathcal{U}$ is the uniform distribution, and "$\pm$" means with equal chance.) Then, we proceed to calculate $J^1(x; \varphi_1, \varphi_2^*)$ by Monte Carlo simulation as before. By definition of the Nash equilibrium, $J^1(x; \varphi_1, \varphi_2^*)$ can not be larger than $V_1(x)$. Within numerical tolerance, this is indeed observed in Figure 4, where the blue empty circles (representing $J^1(x; \varphi_1, \varphi_2^*)$) do not lie over the blue solid curve (which represents $V_1(x)$). Full blue circles represent $J^2(x; \varphi_1, \varphi_2^*)$: note that the player who sticks to her optimal strategy may indeed improve over $V_2(x)$,

should her opponent depart from a Nash equilibrium. (When player 2 is the one who changes, the red circles and red curve in Figure 4 apply instead.)

We stress, however, that Monte Carlo simulations such as those cannot *prove* that a pair of strategies form a Nash equilibrium. (At most, they could disprove it.) The only way—in the current state of the theory—of fully characterizing a Nash equilibrium calls for solving the system of QVIs—which Algorithm 2 has now made possible.

4.2. Capped benchmark game

In this game, we replace the running payoffs of the benchmark by a version capped at $K > 0$:

$$\bar{f}_i(x) := \min\{(-1)^{i-1}(x - s_i), K\} \tag{13}$$

We shall always take $K = 5$. Once again, the corresponding unilateral games are well-posed and their solutions can be used as initial guess for the capped benchmark game. As in the previous example, the capped game also seems to converge from the zero guess. Some convergence results are compiled in Table 2. Note that convergence falters with $M = 600$ and $M = 3300$ (independently of $k_{\max}$); we will come back to this later.

M	R^∞	its.
600	—	∞
900	6.5×10^{-10}	100
1200	8.8×10^{-9}	124
1500	7.4×10^{-9}	78
2700	4.1×10^{-9}	96
3000	6.7×10^{-9}	125
3300	—	∞

TABLE 2. Convergence of R^∞ in the capped benchmark game using the zero guess. The hyphen stands for lack of convergence within ε. Same parameters as in Figure 5.
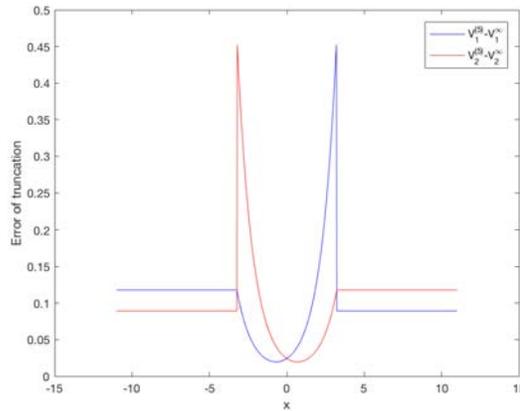

FIGURE 5. Difference (in absolute value, less than 1%) between the (numerical, with $M = 1200$) value functions of the capped benchmark game and those (exact) of the benchmark game—justifying the former being used as initial guess for the latter. Parameters are: $\sigma = 0.25, \rho = 0.03, c = 100, \tilde{c} = 30, \lambda = 0.5, \tilde{\lambda} = 0.3, s_1 = -\pi/3, s_2 = \pi/3$ and $K = 5$ (for capped version of game).

The value functions of the capped benchmark game are a good approximation to those of the benchmark game itself (see Figure 5). This seems to make sense: due to the action of the opponent and for $\sigma \ll K$, the discarded portion of the payoff is not very relevant in practice.

4.3. **Benchmark game with educated initial guess**

Finally, we tackle the benchmark game, for which an exact solution is available (see Section 1.3). Contrary to the previous examples, Algorithm 2 does not seem to enjoy unconditional convergence here. In fact, when the zero guess was used, it failed to converge more often than not (not reported). In order to construct an adequate initial guess, we first solve for the value functions of the capped benchmark game. Using them as the initial guess, convergence was achieved in every experiment.
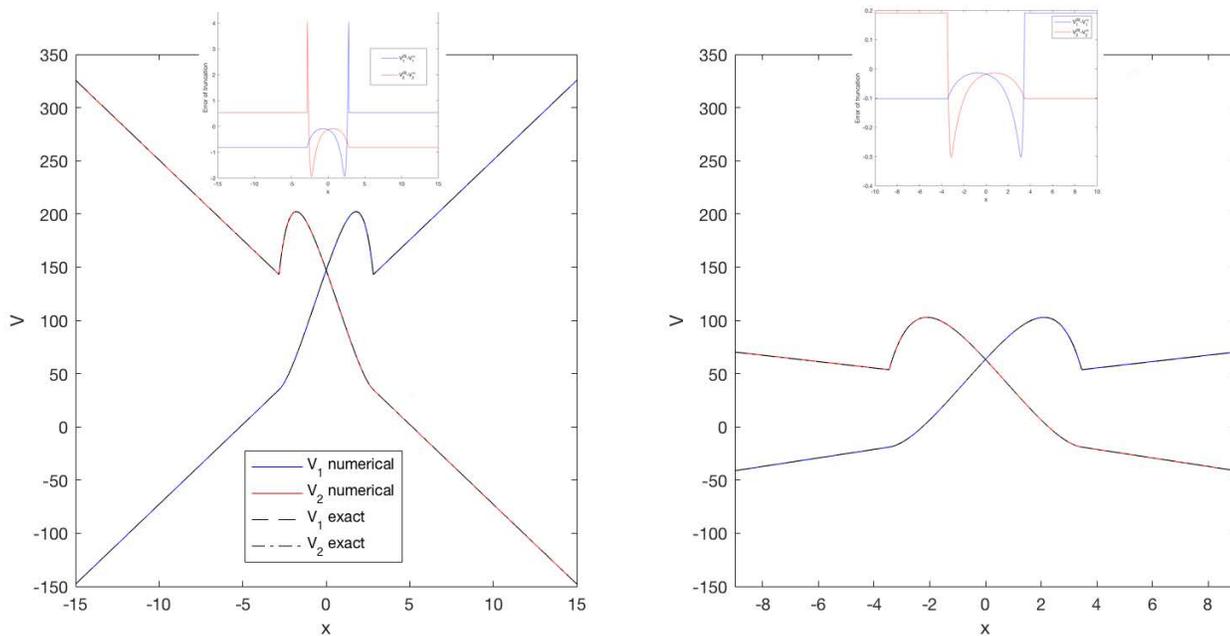

FIGURE 6. Value functions and Nash equilibrium of two instances of benchmark game. Initial guess: solutions of capped benchmark games ($K = 5$). Overlaid, error of the initial guess. Parameters: $\rho = 0.02$, $\sigma = 0.15$, $s_1 = -3$, $s_2 = 3$, $c = 100$, $\tilde{c} = 0$, $\lambda = \tilde{\lambda} = 15$, $M = 1000$ (left); $\rho = 0.03$, $\sigma = 0.25$, $s_1 = -2$, $s_2 = 2$, $c = 100$, $\tilde{c} = 30$, $\lambda = 4$, $\tilde{\lambda} = 3$, $M = 1000$ (right).

The result of two experiments are plotted on Figure 6. The numerical approximations can hardly be distinguished from the exact solutions with the naked eye. As it was done with the parabolic game, once again we can retrieve an approximate Nash equilibrium from the numerical solution. For the left figure, this is $\varphi_1^* = \big((-2.82, +\infty), 1.53 - x\big)$ and $\varphi_2^* = \big((-\infty, 2.82), -1.53 - x\big)$. When compared with the exact equilibrium, the errors on the corresponding abscissae are smaller than 0.15, i.e., smaller than half the grid step.

On Tables 3 and 4, the convergence of the numerical approximation provided by Algorithm 2 to the true solution is demonstrated. However, R^∞ fails to drop below ε for a fine enough discretisation. This reflects a pattern: R^∞ stagnates as M grows. Upon closer inspection, it turns out that the stagnating largest pointwise residual for each player takes place at the junction between the intervention and continuation regions of the opponent, where the exact value function of the former player has a singularity (a *kink*). As discussed in Section 3, Algorithm 2 is always going to place those kinks at finite difference nodes. The overshooting residuals are thus due to the inability of a numerical solution to reproduce a sharp, nonsmooth feature—where the finite

difference derivatives grow unbounded as the distance between nodes goes to zero. Figure 7 illustrates this situation. It can be seen that errors continue to be acceptable (far less than 1% relative error in the worst case). We also highlight the fact that the kinks do not seem to bring about oscillations of the numerical solutions around them—the notorious Gibbs' phenomenon.

M	R^∞	|error|	its.
500	$3.8{\times}10^{-10}$	0.687	126
1000	$1.2{\times}10^{-9}$	0.805	154
1500	$7.2{\times}10^{-9}$	0.512	157
2000	$3.5{\times}10^{-9}$	0.365	172
2500	—	—	∞

TABLE 3. Convergence of Algorithm 2 for benchmark game. Same parameters as in Figure 6 (left).

M	R^∞	|error|	its.
600	—	—	∞
800	$9.5{\times}10^{-10}$	0.023	183
1000	$3.7{\times}10^{-9}$	0.330	177
1400	$8.7{\times}10^{-9}$	0.196	159
1800	$6.8{\times}10^{-9}$	0.121	226
2200	$5.6{\times}10^{-9}$	0.073	224
2600	—	—	∞

TABLE 4. Convergence of Algorithm 2 for benchmark game (same parameters as in Figure 6 (right).

5. CONCLUSIONS

We have designed and tested a novel policy iteration algorithm—the first one as far as we know—to numerically solve nonzero-sum stochastic impulse games (NZSSIGs). The approach consists in solving a system of quasi-variational inequalities which characterizes the value functions and Nash equilibrium, exploiting a recent theoretical breakthrough in [1].

Our algorithm computes iteratively the approximate solution by partitioning, for each of the players, the discretised spatial domain into an approximate continuation region and an approximate intervention region. They are defined through a relaxation parameter that evolves along the iterations. In the continuation region, we solve one quasi-variational inequality by means of (a generalization of) Howard's algorithm, whereas in the complement, a gain is computed. A strategy for producing an educated initial guess to start the iterations—which relies on solving two associated, standard impulse control problems—has been presented along with the new Algorithm as well.

We have not carried out a convergence analysis. Instead, we have gathered plenty of numerical evidence showing that the Algorithm can be applied confidently. In particular, we have performed convergence tests on the only available—at the time of writing—example of an analytically solvable NZSSIG. In other test NZSSIGs for which we do not have an analytic solution, we have explored consistency and numerical results on the value functions and Nash equilibrium, also with satisfactory results.

Value functions of NZSSIGs can develop sharp kinks at the confluence between the continuation and the intervention regions of the opponent. Capturing such features into numerical approximations is a pervasive challenge of numerical analysis. Numerical tests show that the presence of such kinks may eventually put a cap
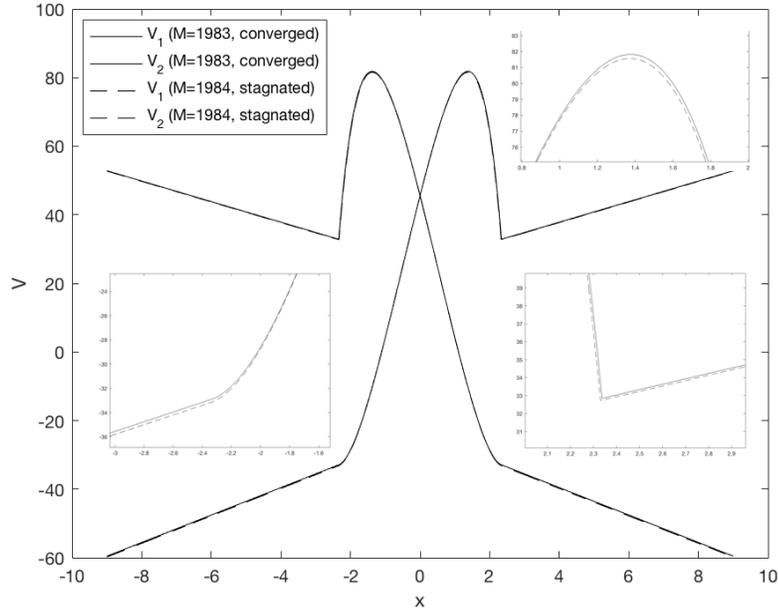
FIGURE 7. Stagnation of accuracy: the pair with $M = 1983$ is fully converged (within numerical tolerance); the pair with $M = 1984$ (overlaid) is not. The insets zoom in on both pairs of functions. Parameters: $\rho = 0.02$, $\sigma = 0.15$, $s_1 = -1$, $s_2 = 1$, $c = 100$, $\tilde{c} = 30$, $\lambda = 4$, $\tilde{\lambda} = 3$.

to the accuracy attained by our algorithm. On the other hand, its stability is not affected by them. Moreover, in every case the largest pointwise error was perfectly acceptable for the purposes of most applications.

In sum, the new Algorithm offers a means of gaining quantitative insight into applications modelled by NZSSIGs. Natural continuations of the present work include the convergence analysis of the Algorithm, and enriching the discretisation method so as to better capture singularities in the solutions.

ACKNOWLEDGMENTS

REFERENCES

[1] R. Aïd, M. Basei, G. Callegaro, L. Campi, T. Vargiolu, *Nonzero-sum stochastic differential games with impulse controls: a verification theorem with applications*, arxiv.org/abs/1605.00039 (2017).

[2] M. Basei, *Topics in stochastic control and differential game theory, with application to mathematical finance*. Ph.D. Thesis in Mathematics, University of Padova (2016).

[3] R. Bellman, *Functional equations in the theory of dynamic programming-XVI: An equation involving iterates*. Journal of Mathematical Analysis and Applications **30**(3), 592-595 (1970).

[4] R. Bellman, *Dynamic Programming. With a new introduction by Stuart Dreyfus. Reprint of the 1957 edition*. Princeton Landmarks in Mathematics, Princeton University Press, Princeton (2010).

[5] O. Bokanowski, S. Maroso and H. Zidani, *Some convergence results for Howard's algorithm*, SIAM J. Numer. Anal., **47**(4), 3001-3026 (2009).

[6] A. Cosso, *Stochastic Differential Games Involving Impulse Controls and Double-Obstacle Quasi-Variational Inequalities*. SIAM Journal on Control and Optimization **51**(3), 2102-2131 (2013).

[7] J.P. Chancelier, M. Messaoud, A. Sulem, *A policy iteration algorithm for fixed point problems with nonexpansive operators.* Mathematical Methods of Operations Research **65**(2), 239-259 (2007).

[8] J.P. Chancelier, B. Oksendal, A. Sulem, *Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control.* Tr. Mat. Inst. Steklova **237**, 149-172 (2002).

[9] R.A. Howard, *Dynamic Programming and Markov Processes.* Technology Press and Wiley, New York (1960).

[10] B. Øksendal and A. Sulem, *Applied stochastic control of jump diffusions.* 2nd Edition, Springer, Berlin (2007).