



# A Fixed-Point Policy-Iteration-Type Algorithm for Symmetric Nonzero-Sum Stochastic Impulse Control Games

Diego Zabaljauregui<sup>1</sup> 

© The Author(s) 2020

## Abstract

Nonzero-sum stochastic differential games with impulse controls offer a realistic and far-reaching modelling framework for applications within finance, energy markets, and other areas, but the difficulty in solving such problems has hindered their proliferation. Semi-analytical approaches make strong assumptions pertaining to very particular cases. To the author's best knowledge, the only numerical method in the literature is the heuristic one we put forward in Aid et al (ESAIM Proc Surv 65:27–45, 2019) to solve an underlying system of quasi-variational inequalities. Focusing on symmetric games, this paper presents a simpler, more precise and efficient fixed-point policy-iteration-type algorithm which removes the strong dependence on the initial guess and the relaxation scheme of the previous method. A rigorous convergence analysis is undertaken with natural assumptions on the players strategies, which admit graph-theoretic interpretations in the context of weakly chained diagonally dominant matrices. A novel provably convergent single-player impulse control solver is also provided. The main algorithm is used to compute with high precision equilibrium payoffs and Nash equilibria of otherwise very challenging problems, and even some which go beyond the scope of the currently available theory.

**Keywords** Stochastic differential games · Nonzero-sum games · Impulse control · Nash equilibrium · Quasi-variational inequality · Howard's algorithm · Fixed point policy iteration · Weakly chained diagonally dominant matrix

## Introduction

Stochastic differential games model the interaction between players whose objective functions depend on the evolution of a certain continuous-time stochastic process. The subclass of impulse games focuses on the case where the players only act at discrete

---

✉ Diego Zabaljauregui  
d.zabaljauregui@lse.ac.uk

<sup>1</sup> Department of Statistics, London School of Economics and Political Science, London, UK

(usually random) points in time by shifting the process. In doing so, each of them incurs into costs and possibly generates “gains” for the others at the same time. They constitute a generalization of the well-known (single-player) optimal impulse control problems [33, Chpt.7-10], which have found a wide range of applications in finance, energy markets and insurance [5,9,13,23,31], among plenty of other fields.

From a deterministic numerical viewpoint, an impulse control problem entails the resolution of a differential quasi-variational inequality (QVI) to compute the value function and, when possible, retrieve an optimal strategy. Policy-iteration-type algorithms [4,15,16] undoubtedly occupy an ubiquitous place in this respect, especially in the infinite horizon case.

The presence of a second player makes matters much more challenging, as one needs to find two optimal (or *equilibrium*) payoffs dependent on one another, and the optimal strategies take the form of Nash equilibria (NEs). And while impulse controls give a more realistic setting than “continuous” controls in applications such as the aforementioned, they normally lead to less tractable and technical models.

It is not surprising then, that the literature in impulse games is limited and mainly focused on the zero-sum case [8,17,21]. The more general and versatile nonzero-sum instance has only recently received attention. The authors of [1] consider for the first time a general two-player game where both participants act through impulse controls,<sup>1</sup> and characterize certain type of equilibrium payoffs and NEs via a system of QVIs by means of a verification theorem. Using this result, they provide the first example of an (almost) fully analytically solvable game, motivated by central banks competing over the exchange rate. The result is generalized to  $N$  players in [10], which also gives a semi-analytical solution (i.e., depending on several parameters found numerically) to a concrete cash management problem.<sup>2</sup> A different, more probabilistic, approach is taken in [24] to find a semi-analytical solution to a strategic pollution control problem and to prove another verification theorem.

The previous examples, and the lack of others,<sup>3</sup> give testimony of how difficult it is to explicitly solve nonzero-sum impulse games. The analytical approaches require an educated guess to start with and (with the exception of the *linear game* in [1]) several parameters need to be solved for in general from highly-nonlinear systems of equations coupled with order conditions. All of this can be very difficult, if not prohibitive, when the structure of the game is not simple enough. Further, all of them (as well as the majority of concrete examples in the impulse control literature) assume linear costs. In general, for nonlinear costs, the state to which each player wants to shift the process when intervening is not unique and it depends on the starting point. This effectively means that infinite parameters may need to be solved for, drastically discouraging this methodology.

While the need for numerical schemes able to handle nonzero-sum impulse games is obvious, unlike in the single-player case, this is an utterly underdeveloped line of

<sup>1</sup> [18,39] also consider nonzero-sum impulse games but assume the intervention times of the players are known from the outset.

<sup>2</sup> [10] also studies the mean field limit game.

<sup>3</sup> [1] also gives semi-analytical solutions to modifications of the linear game when changing the payoffs in a non-symmetric way. To the best of the author’s knowledge, these are the only examples available at the time of writing.

research. Focusing on the purely deterministic approach, solving the system of QVIs derived in [1] involves handling coupled free boundary problems, further complicated by the presence of nonlinear, nonlocal and noncontractive operators. Additionally, solutions will typically be irregular even in the simplest cases such as the linear game. Moreover, the absence of a viscosity solutions framework such as that of impulse control [37] means that it is not possible to know whether the system of QVIs has a solution (not to mention some form of uniqueness) unless one can explicitly solve it. This is further exacerbated by the fact that even defining such a system requires a priori assumptions on the solution (the *unique impulse property*). This is also the case in [24].

To the author's best knowledge, the only numerical method available in the literature is our algorithm in [3], which tackles the system of QVIs by sequentially solving single-player impulse control problems combined with a relaxation scheme. Unfortunately, the choice of the relaxation scheme is not obvious in general and the convergence of the algorithm relies on a good initial guess. It was also observed that stagnation could put a cap on the accuracy of the results, without any simple solution to it. Lastly, while numerical validation was performed, no rigorous convergence analysis was provided.

Restricting attention to the one-dimensional infinite horizon two-player case, this paper puts the focus on certain nonzero-sum impulse games which display a symmetric structure between the players. This class is broad enough to include many interesting applications; no less than the competing central banks problem (whether in its linear form [1] or others considered in the single bank formulation [4,19,30,32]), the cash management problem [10] (reducing its dimension by a simple change of variables) and the generalization of many impulse control problems to the two-player case.

For this class of games, an iterative algorithm is presented which substantially improves [3, Alg.2] by harnessing the symmetry of the problem, removing the strong dependence on the initial guess and dispensing with the relaxation scheme altogether. The result is a simpler and more intuitive, precise and efficient routine, for which a convergence analysis is provided. It is shown that the overall routine admits a representation that strongly resembles, both algorithmically and in its properties, that of the combined fixed-point policy-iteration methods [14,26], albeit with nonexpansive operators. Still, a certain contraction property can still be established.

To perform the analysis, we impose assumptions on the discretization scheme used on the system of QVIs and the discrete admissible strategies. These naturally generalize those of the impulse control case [4] and admit graph-theoretic interpretations in terms of weakly chained diagonally dominant (WCDD) matrices and their recently introduced matrix sequences counterpart [7]. We establish a clear parallel between these discrete type assumptions, the behaviour of the players and the Verification Theorem.

Section 1 deals with the analytical problem. Starting with an overview of the model (Sect. 1.1), we recall the Verification Theorem of [1] and the system of QVIs we want to solve (Sect. 1.2). We then give a precise definition of the class of symmetric nonzero-sum impulse games and establish some preliminary results (Sect. 1.3).

Section 2 considers the analogous discrete problem. Section 2.1 specifies a general discrete version of the system of QVIs, such that any discretization scheme compliant with the assumptions to be imposed will enjoy the same properties. Section 2.2

presents the iterative algorithm, and shows how the impulse control problems that need to be sequentially solved have a unique solution that can be handled by policy iteration. Additionally, Sect. 2.3 provides an alternative general solver for impulse control problems. It consists of an instance of fixed-point policy-iteration that is non-compliant with the standard assumptions [26] and, as far as the author knows, was not used in the context of impulse control before, other than heuristically in [3]. We prove its convergence under the present framework.

Section 2.4 characterizes the overall iterative algorithm as a fixed-point policy-iteration-type method, allowing for reformulations of the original problem and results pertaining to the solutions. The necessary matrix and graph-theoretic definitions and results needed are collected in Appendix A for the reader's convenience. Section 2.5 carries on with the overall convergence analysis and shows to which extent different sets of reasonable assumptions are enough to guarantee convergence to solutions, convergence of strategies and boundedness of iterates. Sufficient conditions for convergence are proved. Discretization schemes are provided in Sect. 2.6.

Section 3 presents the numerical results. In Sect. 3.1, a variety of symmetric nonzero-sum impulse games, many seemingly too complicated to be handled analytically, are explicitly solved for equilibrium payoffs and NE strategies with great precision. This is done on a fixed grid, while considering different performance metrics and addressing practical matters of implementation. In the absence of a viscosity solutions framework to establish convergence to analytical solutions as the grid is refined, Sect. 3.2 performs a numerical validation using the only examples of symmetric solvable games in the literature. Section 3.3 addresses the case of games without NEs. Section 3.4 tackles games beyond the scope of the currently available theory, displaying discontinuous impulses and very irregular payoffs. The latter give insight and motivate further research into this field.

## 1 Analytical Continuous-Space Problem

In this section we start by reviewing a general formulation of two-player nonzero-sum stochastic differential games with impulse controls, as considered in [1], together with the main theoretical result of the authors: a characterization of certain NEs via a deterministic system of QVIs. The indices of the players are denoted  $i = 1, 2$ . We will generally use  $i$  to indicate a given player and  $j$  to indicate their opponent. Since no other type of games is considered in this paper, we will often speak simply of “games” for brevity. Afterwards we shall specialize the discussion in the (yet to be specified) symmetric instance.

Throughout the paper, we restrict our attention to the one-dimensional infinite-horizon case. (A similar review is carried out in [3, Sect.1].) Some of the most technical details concerning the well-posedness of the model are left out for brevity and can be found in [1, Sect.2].

### 1.1 General Two-Player Nonzero-Sum Impulse Games

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  be a filtered probability space under the usual conditions supporting a standard one-dimensional Wiener process  $W$ . We consider two players that observe the evolution of a *state variable*  $X$ , modifying it when convenient through *controls* of the form  $u_i = \{(\tau_i^k, \delta_i^k)\}_{k=1}^\infty$  for  $i = 1, 2$ . The stopping times  $(\tau_i^k)$  are their *intervention times* and the  $\mathcal{F}_{\tau_i^k}$ -measurable random variables  $(\delta_i^k)$  are their *intervention impulses*. Given controls  $(u_1, u_2)$  and a starting point  $X_{0-} = x \in \mathbb{R}$ , we assume  $X = X^{x;u_1,u_2}$  has dynamics

$$X_t = x + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s + \sum_{k: \tau_1^k \leq t} \delta_1^k + \sum_{k: \tau_2^k \leq t} \delta_2^k, \tag{1.1.1}$$

for some given drift and volatility functions  $\mu, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ , locally Lipschitz with linear growth.<sup>4</sup>

Equation (1.1.1) states that  $X$  evolves as an Itô diffusion in between the intervention times, and that each intervention consists in shifting  $X$  by applying an impulse. It is assumed that the players choose their controls by means of threshold-type *strategies* of the form  $\varphi_i = (\mathcal{I}_i, \delta_i)$ , where  $\mathcal{I}_i \subseteq \mathbb{R}$  is a closed set called *intervention (or action) region* and  $\delta_i : \mathbb{R} \rightarrow \mathbb{R}$  is an *impulse function* assumed to be continuous. The complement  $\mathcal{C}_i = \mathcal{I}_i^c$  is called *continuation (or waiting) region*.<sup>5</sup> That is, player  $i$  intervenes if and only if the state variable reaches her intervention region, by applying an impulse  $\delta_i(X_{t-})$  (or equivalently, shifting  $X_{t-}$  to  $X_{t-} + \delta_i(X_{t-})$ ). Further, we impose a priori constraints on the impulses: for each  $x \in \mathbb{R}$  there exists a set  $\emptyset \neq \mathcal{Z}_i(x) \subseteq \mathbb{R}$  (further specified in Sect. 1.3) such that  $\delta_i(x) \in \mathcal{Z}_i(x)$  if  $x \in \mathcal{I}_i$ .<sup>6</sup> We also assume the game has no end and player 1 has the priority should they both want to intervene at the same time. (The latter will be excluded later on; see Definition 1.3.1 and the remarks that follow it.)

Given a starting point and a pair strategies, the (expected) *payoff* of player  $i$  is given by

$$J_i(x; \varphi_1, \varphi_2) := \mathbb{E} \left[ \int_0^\infty e^{-\rho_i s} f_i(X_s)ds - \sum_{k=1}^\infty e^{-\rho_i \tau_i^k} c_i(X_{(\tau_i^k)^-}, \delta_i^k) + \sum_{k=1}^\infty e^{-\rho_i \tau_j^k} g_i(X_{(\tau_j^k)^-}, \delta_j^k) \right],$$

with  $X = X^{x;u_1,u_2} = X^{x;\varphi_1,\varphi_2}$ . For player  $i$ ,  $\rho_i > 0$  represents her (subjective) *discount rate*,  $f_i : \mathbb{R} \rightarrow \mathbb{R}$  her *running payoff*,  $c_i : \mathbb{R}^2 \rightarrow (0, +\infty)$  her *cost of*

<sup>4</sup> See [1, Def.2.2] for a precise recursive definition in terms of the strategies.

<sup>5</sup> In [1], strategies are described in terms of continuation regions instead.

<sup>6</sup> In [1],  $\mathcal{Z}_i(x)$  is the same for every  $x \in \mathbb{R}$ . The generalization in this paper is a standard one in impulse control and will prove useful in the sequel. The results in [1] still hold with the same proofs.

intervention and  $g_i : \mathbb{R}^2 \rightarrow \mathbb{R}$  her gain due to her opponent's intervention (not necessarily non-negative). The functions  $f_i, c_i, g_i$  are assumed to be continuous.

Throughout the paper, only admissible strategies are considered. Briefly,  $(\varphi_1, \varphi_2)$  is *admissible* if it gives well-defined payoffs for all  $x \in \mathbb{R}$ ,  $\|X\|_\infty$  has finite moments and, although each player can intervene immediately after the other, infinite simultaneous interventions are precluded.<sup>7</sup> As an example, if the running payoffs have polynomial growth, the “never intervene strategies”  $\varphi_1 = \varphi_2 = (\emptyset, \emptyset \hookrightarrow \mathbb{R})$  are admissible and the game can be played.

Given a game, we want to know whether it admits some Nash equilibrium and how to compute it. Recall that a pair of strategies  $(\varphi_1^*, \varphi_2^*)$  is a *Nash equilibrium* (NE) if for every admissible  $(\varphi_1, \varphi_2)$ ,

$$J_1(x; \varphi_1^*, \varphi_2^*) \geq J_1(x; \varphi_1, \varphi_2^*) \quad \text{and} \quad J_2(x; \varphi_1^*, \varphi_2^*) \geq J_2(x; \varphi_1^*, \varphi_2),$$

i.e., no player can gain from a unilateral change of strategy. If one such NE exists, we refer to  $(V_1, V_2)$ , with  $V_i(x) = J_i(x; \varphi_1^*, \varphi_2^*)$ , as a pair of *equilibrium payoffs*.

### 1.2 General System of Quasi-Variational Inequalities

To present the system of QVIs derived in [1], we need to define first the *intervention operators*. For any  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$  and  $x \in \mathbb{R}$ , the *loss operator* of player  $i$  is defined as<sup>8</sup>

$$\mathcal{M}_i V_i(x) := \sup_{\delta \in \mathcal{Z}_i(x)} \{V_i(x + \delta) - c_i(x, \delta)\}. \tag{1.2.1}$$

When applied to an equilibrium payoff, the loss operator  $\mathcal{M}_i$  gives a recomputed present value for player  $i$  due to the cost of her own intervention. Given the optimality of the NEs, one would intuitively expect that  $\mathcal{M}_i V_i \leq V_i$  for equilibrium payoffs and that the equality is attained only when it is optimal for player  $i$  to intervene. Under this logic:

**Definition 1.2.1** We say that the pair  $(V_1, V_2)$  has the *unique impulse property* (UIP) if for each  $i = 1, 2$  and  $x \in \{\mathcal{M}_i V_i = V_i\}$ , there exists a unique impulse, denoted  $\delta_i^*(x) = \delta_i^*(V_i)(x) \in \mathcal{Z}_i(x)$ , that realizes the supremum in (1.2.1).<sup>9</sup>

If  $(V_1, V_2)$  enjoys the UIP, we define the *gain operator* of player  $i$  as

$$\mathcal{H}_i V_i(x) := V_i(x + \delta_j^*(x)) + g_i(x, \delta_j^*(x)), \quad \text{for } x \in \{\mathcal{M}_j V_j = V_j\} \tag{1.2.2}$$

When applied to equilibrium payoffs, the gain operator  $\mathcal{H}_i$  gives a recomputed present value for player  $i$  due to her opponent's intervention.

<sup>7</sup> More precisely, these would be  $\mathbb{R}$ -admissible strategies. See [1, Def.2.5] for more details.

<sup>8</sup> Although we could have  $\mathcal{M}_i V_i(x) = +\infty$ , this will be excluded when enforcing the system of QVIs (1.2.3).

<sup>9</sup> We do not require the UIP to hold outside of  $\{\mathcal{M}_i V_i = V_i\}$ , as this is not the case for equilibrium payoffs in many examples, such as the linear game with constant costs/gains. Proofs in [1] carry through unaltered.

Finally, let us denote by  $\mathcal{A}$  the infinitesimal generator of  $X$  when uncontrolled, i.e.,

$$\mathcal{A}V(x) := \frac{1}{2}\sigma^2(x)V''(x) + \mu(x)V'(x),$$

for any  $V : \mathbb{R} \rightarrow \mathbb{R}$  which is  $C^2$  at some open neighborhood of a given  $x \in \mathbb{R}$ . We assume this regularity holds whenever we compute  $\mathcal{A}V(x)$  for some  $V$  and  $x$ . The following Verification Theorem, due to [1, Thm.3.3], states that if a regular enough solution  $(V_1, V_2)$  to a certain system of QVIs exists, then it must be a pair of equilibrium payoffs, and a corresponding NE can be retrieved. We state here a simplified version that applies to the one-dimensional infinite-horizon games at hand.<sup>10</sup>

**Theorem 1.2.2** (General system of QVIs) *Given a game as in Sect. 1.1, let  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$  be pair of functions with the UIP, such that for any  $i, j \in \{1, 2\}, i \neq j$ :*

$$\begin{cases} \mathcal{M}_j V_j - V_j \leq 0 & \text{on } \mathbb{R} \\ \mathcal{H}_i V_i - V_i = 0 & \text{on } \{\mathcal{M}_j V_j - V_j = 0\} =: \mathcal{I}_j^* \\ \max \{ \mathcal{A}V_i - \rho_i V_i + f_i, \mathcal{M}_i V_i - V_i \} = 0 & \text{on } \{\mathcal{M}_j V_j - V_j < 0\} =: \mathcal{C}_j^* \end{cases} \tag{1.2.3}$$

and  $V_i \in C^2(\mathcal{C}_j^* \setminus \partial \mathcal{C}_i^*) \cap C^1(\mathcal{C}_j^*) \cap C(\mathbb{R})$  has polynomial growth and bounded second derivative on some reduced neighbourhood of  $\partial \mathcal{C}_i^*$ . Suppose further  $(\mathcal{I}_i^*, \delta_i^*)_{i=1,2}$  are admissible strategies.<sup>11</sup>

Then,  $(V_1, V_2)$  are equilibrium payoffs attained at a NE  $(\mathcal{I}_i^*, \delta_i^*)_{i=1,2}$ .

The first equation of system (1.2.3) states that at an equilibrium, a player cannot increase her own payoff by a unilateral intervention. One therefore expects that the equality  $\mathcal{M}_j V_j = V_j$  will only hold when player  $j$  intervenes, or in other words, when the value she gains can compensate the cost of her intervention. Consequently, the second equation says that a gain results from the opponent’s intervention. Finally, the last one, means that when the opponent does not intervene, each player faces a single-player impulse control problem.

We conclude this section with some final observations that will be relevant in the sequel:

**Remark 1.2.3** An immediate consequence of assuming strictly positive costs is that intervening at any state with a null impulse reduces the payoff of the acting player and is therefore suboptimal. This is also displayed in system (1.2.3): if at some state  $x$ ,  $\mathcal{M}_j V_j(x)$  was realized for  $\delta = 0$ , then  $\mathcal{M}_j V_j(x) = V_j(x) + c_j(x, 0) < V_j(x)$ . At the same time, allowing for vanishing costs often leads to degenerate games in the current framework [1, Sect.4.4]. Hence, assuming  $c_j > 0$  is quite reasonable.

<sup>10</sup> Unlike in [1], there is no terminal condition in the system of QVIs and the assumption that  $\partial \mathcal{C}_i^*$  be a Lipschitz surface is trivially satisfied for an open  $\mathcal{C}_i^* \subseteq \mathbb{R}$ , as it is a countable union of disjoint open intervals.

<sup>11</sup> For consistency with the strategies’ definition, one should assume that  $\delta^*$  has been continuously extended to  $\mathbb{R}$ . The conclusion is unaffected by the choice of the extension.

**Remark 1.2.4** Consider the case of nonnegative impulses and cost functions being strictly concave in the impulse as in [17,21]. That is,  $c_i(x, \delta + \bar{\delta}) < c_i(x, \delta) + c_i(x + \delta, \bar{\delta})$  for all  $x \in \mathbb{R}$ ,  $\delta, \bar{\delta} \geq 0$ . This models the situation in which simultaneous interventions are more expensive than a single one to the same effect. In such cases, it is easy to see that in the context of Theorem 1.2.2, player  $i$  will only shift the state variable towards her continuation region.<sup>12</sup>

### 1.3 Symmetric Two-Player Nonzero-Sum Impulse Games

We want to focus our study on games which present a certain type of symmetric structure between the players, generalising the linear game [1] and the cash management game [10].<sup>13</sup>

**Notation** The type of games presented in Sect. 1.1 are fully defined by setting the drift, volatility, impulse constraints, discount rates, running payoffs, costs and gains. In other words, any such game can be represented by a tuple  $\mathcal{G} = (\mu, \sigma, \mathcal{Z}_i, \rho_i, f_i, c_i, g_i)_{i=1,2}$ .

**Definitions 1.3.1** We say that a game  $\mathcal{G} = (\mu, \sigma, \mathcal{Z}_i, \rho_i, f_i, c_i, g_i)_{i=1,2}$  is *symmetric (with respect to zero)* if

- (S1)  $\mu$  is odd and  $\sigma$  is even (i.e.,  $\mu(x) = -\mu(-x)$  and  $\sigma(x) = \sigma(-x)$  for all  $x \in \mathbb{R}$ ).
- (S2)  $-\mathcal{Z}_2(-x) = \mathcal{Z}_1(x) \subseteq [0, +\infty)$  for all  $x \in \mathbb{R}$  and  $\mathcal{Z}_1(x) = \{0\} = \mathcal{Z}_2(-x)$  for all  $x \geq 0$ .
- (S3)  $\rho_1 = \rho_2$ ,  $f_1(x) = f_2(-x)$ ,  $c_1(x, \delta) = c_2(-x, -\delta)$  and  $g_1(x, -\delta) = g_2(-x, \delta)$ , for all  $\delta \in \mathcal{Z}_1(x)$ ,  $x \in \mathbb{R}$ .

We say that the game is *symmetric with respect to  $s$*  (for some  $s \in \mathbb{R}$ ), if the  $s$ -shifted game  $(\mu(x + s), \sigma(x + s), \mathcal{Z}_i(x + s), \rho_i, f_i(x + s), c_i(x + s, \delta), g_i(x + s, \delta))_{i=1,2}$  is symmetric. We refer to  $x = s$  as a *symmetry line* of the game.

Condition (S1) is necessary for the state variable to have symmetric dynamics. In particular, together with (S3), it guarantees symmetry between solutions of the Hamilton–Jacobi–Bellman (HJB) equations of the players when there are no interventions, i.e.,

$$\begin{aligned}
 V &= V^*(x) \text{ solves } \mathcal{A}V - \rho_1 V + f_1 = 0 \quad \text{if and only if} \\
 V &= V^*(-x) \text{ solves } \mathcal{A}V - \rho_2 V + f_2 = 0.
 \end{aligned}$$

**Examples 1.3.2** The most common examples of Itô diffusions satisfying this assumption are the scaled Brownian motion (symmetric with respect to zero) and the Ornstein–Uhlenbeck (OU) process (symmetric with respect to its long term mean).

<sup>12</sup> Let  $x \in \mathcal{I}_i^*$  and suppose  $y_i^* := x + \delta_i^*(x) \in \mathcal{I}_i^*$ . Set  $y_i^{**} := y_i^* + \delta_i^*(y_i^*)$ . Then, by the UIP, the definitions of  $\delta_i^*(x)$  and  $\mathcal{I}_i^*$ , and the concavity of the cost:  $V_i(y_i^{**}) - c_i(x, \delta_i^*(x) + \delta_i^*(y_i^*)) \leq V_i(y_i^*) - c_i(x, \delta_i^*(x)) = V_i(y_i^{**}) - c_i(y_i^*, \delta_i^*(y_i^*)) - c_i(x, \delta_i^*(x)) < V_i(y_i^{**}) - c_i(x, \delta_i^*(x) + \delta_i^*(y_i^*))$ , which is a contradiction.

<sup>13</sup> The latter can be reduced to one dimension with the change of variables  $x = x_1 - x_2$ . Additionally, we will restrict attention to unidirectional impulses, as these yield the “most relevant” NE found in [10].

Condition (S3) is self-explanatory, while (S2) is only partly so. Indeed, although symmetric constraints on the impulses  $\mathcal{Z}_1(x) = -\mathcal{Z}_2(-x)$  should clearly be a requirement, the rest of (ii) is in fact motivated by the numerical method to be presented and the type of problems it can handle. On the one hand, the third equation of the QVIs system (1.2.3) implies that a stochastic impulse control problem for player  $i$  needs to be solved on  $\mathcal{C}_j^*$ . The unidirectional impulses assumption is a common one for the convergence of policy iteration algorithms in impulse control.<sup>14</sup> However, it is often too restrictive for many interesting applications,<sup>15</sup> such as when the controller would benefit the most from keeping the state variable within some bounded interval instead of simply keeping it “high” or “low” (see, e.g., [9] and [4, Sect.6.1]). Interestingly enough, assuming unidirectional impulses turns out to be less restrictive when there is a second player present, with an opposed objective. Indeed, it can happen that each player needs not to intervene in one of the two directions, and can instead rely on her opponent doing so, while capitalising a gain rather than paying a cost. See examples in Sect. 3.1 with quadratic and degree four running payoffs.

On the other hand,  $\mathcal{Z}_1(x) = \{0\} = \mathcal{Z}_2(-x)$  for all  $x \geq 0$  means that we can assume without loss of generality that the admissible intervention regions do not cross over the symmetry line; i.e.,  $\mathcal{I}_1 \subseteq (-\infty, 0)$  and  $\mathcal{I}_2 \subseteq (0, +\infty)$  for every pair of strategies. (See Remark 1.2.3.) This guarantees in particular that the players never want to intervene at the same time and the priority rule can be disregarded.

There are different reasons why the last mentioned condition is less restrictive than it first appears to be. It is not uncommon to assume connectedness of either intervention or continuation regions (or other conditions implying them) both in impulse control [22] and nonzero-sum games [20, Sect.1.2.1]. The same can be said for assumptions that prevent the players from intervening in unison [20, Sect.1.2.1],[17, Rmk.6.5].<sup>16</sup> In the context of symmetric games and payoffs (see Lemma 1.3.7) such assumptions would necessarily imply the intervention regions need to be on opposed sides of the symmetry line. Additionally, without any further requirements, strategies such that  $\mathcal{I}_1 \supseteq (-\infty, 0]$  and  $\mathcal{I}_2 \supseteq [0, +\infty)$  would be inadmissible in the present framework, as per yielding infinite simultaneous impulses.

**Definitions 1.3.3** Given a symmetric game, we say that  $((\mathcal{I}_i, \delta_i))_{i=1,2}$  are *symmetric strategies (with respect to zero)* if  $\mathcal{I}_1 = -\mathcal{I}_2$  and  $\delta_1(x) = -\delta_2(-x)$ . Given a symmetric game with respect to some  $s \in \mathbb{R}$ , we say that  $((\mathcal{I}_i, \delta_i))_{i=1,2}$  are *symmetric strategies with respect to  $s$*  if  $((\mathcal{I}_i - s, \delta_i(x + s)))_{i=1,2}$  are symmetric, and we refer to  $x = s$  as a *symmetry line* of the strategies.

**Definition 1.3.4** We say that  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$  are *symmetric functions (with respect to zero)* if  $V_1(x) = V_2(-x)$ . We say that they are *symmetric functions with respect to*

<sup>14</sup> See this assumption in [15, Sect.4] or [33, Sect.10.4.2], its graph-theoretic counterpart in [4, Asm.(H2) and Thm.4.3], and a counterexample of convergence in its absence in [4, Ex.4.9].

<sup>15</sup> See [34,35] for a novel, globally convergent, combined policy iteration and penalization scheme for general impulse control HJBQVIs.

<sup>16</sup> In [17, Rmk.6.5], assumptions are given to guarantee the zero-sum analogous to  $\mathcal{M}_i V_i \leq \mathcal{H}_i V_i$ , with a strict inequality if such assumptions are slightly strengthened. These inequalities, in the context of Theorem 1.2.2, imply that the equilibrium intervention regions cannot intersect.

$s$  (for some  $s \in \mathbb{R}$ ) if  $V_1(x + s), V_2(x + s)$  are symmetric, and we refer to  $x = s$  as a *symmetry line* for  $V_1, V_2$ .

**Remark 1.3.5** Definition 1.3.3 singles out strategies that share the same symmetry line with the game. For the linear game, for example, the authors find infinitely many NEs [1, Prop.4.7], each presenting symmetry with respect to some point  $s$ , but only one for  $s = 0$  (hence, symmetric in the sense of Definition 1.3.3). At the same time, the latter is the only one for which the corresponding equilibrium payoffs  $V_1, V_2$  have a symmetry line as per Definition 1.3.4. The same is true for the cash management game [10].

**Remark 1.3.6** Throughout the paper we will work only with games symmetric with respect to zero, to simplify the notation. Working with any other symmetry line amounts simply to shifting the game and results back and forth.

**Lemma 1.3.7** For any symmetric game, strategies  $(\varphi_1, \varphi_2)$  and functions  $V_1, V_2 : \mathbb{R} \rightarrow \mathbb{R}$ :

- (i) If  $V_1, V_2$  are symmetric, then  $\mathcal{M}_1 V_1, \mathcal{M}_2 V_2$  are symmetric.
- (ii) If  $V_1, V_2$  are symmetric and have the UIP, then  $\delta_1^*(x) = -\delta_2^*(-x)$  and  $\mathcal{H}_1 V_1, \mathcal{H}_2 V_2$  are symmetric.
- (iii) If  $(\varphi_1, \varphi_2)$  are symmetric, then  $J_1(\cdot; \varphi_1, \varphi_2), J_2(\cdot; \varphi_1, \varphi_2)$  are symmetric.
- (iv) If  $V_1, V_2$  are as in Theorem 1.2.2 and  $(\varphi_1^*, \varphi_2^*)$  is the corresponding NE of the theorem, then  $(\varphi_1^*, \varphi_2^*)$  are symmetric if and only if  $V_1, V_2$  are symmetric.

**Proof** (i) and (ii) are straightforward from the definitions.

To see (iii), one can check with the recursive definition of the state variable [1, Def.2.2] that  $X^{-x; \varphi_1, \varphi_2}$  has the same law as  $-X^{x; \varphi_1, \varphi_2}$  (recall that the continuation regions are simply disjoint unions of open intervals). Noting also that intervention times and impulses are nothing but jump times and sizes of  $X$ , one concludes that  $J_1(x; \varphi_1, \varphi_2) = J_2(-x; \varphi_1, \varphi_2)$ , as intended.

Finally, (iv) is a consequence of (i), (ii) and (iii). □

**Convention 1.3.8** In light of Lemma 1.3.7, for any symmetric game we will often lose the player index from the notations and refer always to quantities corresponding to player 1,<sup>17</sup> henceforth addressed simply as “the player”. Player 2 shall be referred to as “the opponent”. Statements like “ $V$  has the UIP” or “ $V$  is a symmetric equilibrium payoff” are understood to refer to  $(V(x), V(-x))$ . Likewise, “ $(\mathcal{I}, \delta)$  is admissible” or “ $(\mathcal{I}, \delta)$  is a NE” refer to the pair  $(\mathcal{I}, \delta(x)), (-\mathcal{I}, -\delta(-x))$ .

Due to their general lack of uniqueness, it is customary in game theory to restrict attention to specific type of NEs, depending on the problem at hand (see for instance [29] for a treatment within the classical theory). Motivated by Lemma 1.3.7 (iii) and (iv), and by Remark 1.3.5, one can arguably state that symmetric NEs are the most meaningful for symmetric games. Furthermore, Lemma 1.3.7 implies that for symmetric games, one can considerably reduce the complexity of the full system of QVIs (1.2.3) provided the conjectured NE (or equivalently, the pair of payoffs) is symmetric. Using Convention 1.3.8, Theorem 1.2.2 and Lemma 1.3.7 give:

<sup>17</sup> Note that  $g$  will denote  $g(x, \delta) := g_1(x, -\delta)$ , as  $\delta \geq 0$  for player 1, yet  $g_1$  depends on the (negative) impulse of player 2.

**Corollary 1.3.9** (Symmetric system of QVIs) *Given a symmetric game as in Definition 1.3.1, let  $V : \mathbb{R} \rightarrow \mathbb{R}$  be a function with the UIP, such that:*

$$\begin{cases} \mathcal{H}V - V = 0 & \text{on } -\{\mathcal{M}V - V = 0\} =: -\mathcal{I}^* \\ \max\{\mathcal{A}V - \rho V + f, \mathcal{M}V - V\} = 0 & \text{on } -\{\mathcal{M}V - V < 0\} =: -\mathcal{C}^* \end{cases} \tag{1.3.1}$$

and  $V \in C^2(-\mathcal{C}^* \setminus \partial\mathcal{C}^*) \cap C^1(-\mathcal{C}^*) \cap C(\mathbb{R})$  has polynomial growth and bounded second derivative on some reduced neighbourhood of  $\partial\mathcal{C}^*$ . Suppose further that  $(\mathcal{I}^*, \delta^*)$  is an admissible strategy.

Then,  $V$  is a symmetric equilibrium payoff attained at a symmetric NE  $(\mathcal{I}^*, \delta^*)$ .

Note that system (1.3.1) also omits the equation  $\mathcal{M}V - V \leq 0$  as per being redundant. Indeed, by Definition 1.3.1 and Remark 1.2.3, at a NE the player does not intervene above 0, nor the opponent below it. Thus,  $\mathcal{M}V - V \leq \max\{\mathcal{A}V - \rho V + f, \mathcal{M}V - V\} = 0$  on  $-\mathcal{C}^* \supset (-\infty, 0]$  and  $\mathcal{M}V - V < 0$  on  $[0, +\infty)$ .

System (1.3.1) simplifies a numerical problem which is very challenging even in cases of linear structure [3]. In light of the previous, we will focus our attention on symmetric NEs only and numerically solving the reduced system of QVIs (1.3.1).

## 2 Numerical Discrete-Space Problem

In this section we consider a discrete version of the symmetric system of QVIs (1.3.1) over a fixed grid, and propose and study an iterative method to solve it. As it is often done in numerical analysis for stochastic control, for the sake of generality we proceed first in an abstract fashion without making reference to any particular discretization scheme. Instead, we give some general assumptions any such scheme should satisfy for the results presented to hold. Explicit discretization schemes within our framework are presented in Sect. 2.6 and used in Sect. 3.

### 2.1 Discrete System of Quasi-variational Inequalities

From now on we work on a discrete symmetric grid

$$G : x_{-N} = -x_N < \dots < x_{-1} = -x_1 < x_0 = 0 < x_1 < \dots < x_N.$$

$\mathbb{R}^G$  denotes the set of functions  $v : G \rightarrow \mathbb{R}$  and  $S : \mathbb{R}^G \rightarrow \mathbb{R}^G$  denotes the symmetry operator,  $Sv(x) = v(-x)$ . In general, by an ‘‘operator’’ we simply mean some  $F : \mathbb{R}^G \rightarrow \mathbb{R}^G$ , not necessarily linear nor affine unless explicitly stated. We shall identify grid points with indices, functions in  $\mathbb{R}^G$  with vectors and linear operators with matrices; e.g.,  $S = (S_{ij})$  with  $S_{ij} = 1$  if  $x_i = -x_j$  and 0 otherwise. The (partial) order considered in  $\mathbb{R}^G$  and  $\mathbb{R}^{G \times G}$  is the usual pointwise order for functions (elementwise for vectors and matrices), and the same is true for the supremum, maximum and arg-maximum induced by it.

We want to solve the following discrete nonlinear system of QVIs for  $v \in \mathbb{R}^G$ :

$$\begin{cases} H v - v = 0 & \text{on } -\{M v - v = 0\} =: -I^* \\ \max \{L v + f, M v - v\} = 0 & \text{on } -\{M v - v < 0\} =: -C^*, \end{cases} \tag{2.1.1}$$

where  $f \in \mathbb{R}^G$  and  $L : \mathbb{R}^G \rightarrow \mathbb{R}^G$  is a linear operator. The nonlinear operators  $M, H : \mathbb{R}^G \rightarrow \mathbb{R}^G$  are as follows: let  $\emptyset \neq Z(x) \subseteq \mathbb{R}$  be a finite set for each  $x \in G$ , with  $Z(x) = \{0\}$  if  $x \geq 0$ . Set  $Z := \prod_{x \in G} Z(x)$  and for each  $\delta \in Z$  let  $B(\delta) : \mathbb{R}^G \rightarrow \mathbb{R}^G$  be a linear operator,  $c(\delta) \in (0, +\infty)^G$  and  $g(\delta) \in \mathbb{R}^G$ , the three of them being *row-decoupled* in the sense of [4,11] (i.e., row  $x$  of  $B(\delta)$ ,  $c(\delta)$ ,  $g(\delta)$  depends only on  $\delta(x) \in Z(x)$ ). Then

$$M v := \max_{\delta \in Z} \{B(\delta)v - c(\delta)\}, \quad H v = H(\delta^*)v := S B(\delta^*)S v + g(S\delta^*) \tag{2.1.2}$$

$$\text{and } \delta^* = \delta^*(v) := \max_{\delta \in Z} \left( \arg \max \{B(\delta)v - c(\delta)\} \right). \tag{2.1.3}$$

Some remarks are in order. Firstly, in the same fashion as the continuous-space case, the sets  $I^*, C^*$  form a partition of the grid and represent the (discrete) intervention and continuation regions of the player, while  $-I^*, -C^*$  are such regions for the opponent.

Secondly, the general representation of  $M$  follows [4,15]. For the standard choices of  $B(\delta)$ , our definition of  $H$  is the only one for which a discrete version of Lemma 1.3.7 holds true (see Sect. 2.6). However, since  $B$  and  $g$  are row-decoupled,  $S B(\delta^*)S$  and  $g(S\delta^*)$  cannot be, as each row  $x$  depends on  $\delta^*(-x)$ . For this reason and the lack of maximization over  $-I^*$ , there is no obvious way to reduce problem (2.1.1) to a classical Bellman problem:

$$\sup_{\varphi} \{ -A(\varphi)v + b(\varphi) \} = 0, \tag{2.1.4}$$

like in the impulse control case [4], to apply Howard’s policy iteration [11, Ho-1]. Furthermore, unlike in the control case, even with unidirectional impulses and good properties for  $L$  and  $B(\delta)$ , system (2.1.1) may have no solution as in the analytical case [1].

Thirdly, we have defined  $\delta^*$  in (2.1.3) by choosing one particular maximizing impulse for each  $x \in G$ . The main motivation behind fixing one is to have a well defined discrete system of QVIs for every  $v \in \mathbb{R}^G$ . (This is not the case for the analytical problem (1.3.1) where the gain operator  $\mathcal{H}$  is not well defined unless  $V$  has the UIP.) Being able to plug in any  $v$  in (2.1.1) and obtain a residual will be useful in practice, when assessing the convergence of the algorithm (see Sect. 3). Whether a numerical solution verifies, at least approximately, a discrete UIP (and the remaining technical conditions of the Verification Theorem) becomes something to be checked separately a posteriori.

**Remark 2.1.1** Choosing the maximum arg-maximum in (2.1.3) is partly motivated by ensuring a discrete solution will inherit the property of Remark 1.2.4. (The proof remains the same, for the discretizations of Sect. 2.6.) We will also motivate it in terms

of the proposed numerical algorithm in Remark 2.6.4. Note that in [3] the minimum arg-maximum is used instead for both players. Nevertheless, the replication of property (ii), Lemma 1.3.7, dictates that it is only possible to be consistent with [3] for one of the two players (in this case, the opponent).

### 2.2 Iterative Algorithm for Symmetric Games

This section introduces the iterative algorithm developed to solve system (2.1.1), which builds on [3, Alg.2] by harnessing the symmetry of the problem and dispenses with the need for a relaxation scheme altogether. It is presented with a pseudocode that highlights the mimicking of system (2.1.1) and the intuition behind the algorithm; namely:

- The player starts with some suboptimal strategy  $\varphi^0 = (I^0, \delta^0)$  and payoff  $v^0$ , to which the opponent responds symmetrically, resulting in a gain for the player (first equation of (2.1.1); lines 1, 2 and 4 of Algorithm 2.2.1).
- The player improves her strategy by choosing the optimal response, i.e., by solving a single-player impulse control problem through a policy-iteration-type algorithm (second equation of (2.1.1); line 5 of Algorithm 2.2.1).
- This procedure is iterated until reaching a stopping criteria (lines 6–8 of Algorithm 2.2.1).

**Notation 2.2.1** In the following:  $G_{<0}$  and  $G_{\leq 0}$  represent the sets of grid points which are negative and nonpositive respectively, and  $\Phi$  the set of (discrete) strategies

$$\Phi := \{\varphi = (I, \delta) : I \subseteq G_{<0} \text{ and } \delta \in Z\}. \tag{2.2.1}$$

Set complements are taken with respect to the whole grid,  $\text{Id} : \mathbb{R}^G \rightarrow \mathbb{R}^G$  is the identity operator; and given a linear operator  $O : \mathbb{R}^G \rightarrow \mathbb{R}^G \simeq \mathbb{R}^{G \times G}$ ,  $v \in \mathbb{R}^G$  and subsets  $I, J \subseteq G$ ,  $v_I \in \mathbb{R}^I$  denotes the restriction of  $v$  to  $I$  and  $O_{IJ} \in \mathbb{R}^{I \times J}$  the submatrix/operator with rows in  $I$  and columns in  $J$ .

---

#### Algorithm 2.2.1 Iterative algorithm for symmetric games

---

Set  $tol, scale > 0$ .

- 1: Choose initial guess:  $v^0 \in \mathbb{R}^G$
  - 2: Set  $I^0 = \{Lv^0 + f \leq Mv^0 - v^0\} \cap G_{<0}$  and  $\delta^0 = \delta^*(v^0)$
  - 3: **for**  $k = 0, 1, \dots$  **do**
  - 4:  $v^{k+1/2} = \begin{cases} H(\delta^k)v^k & \text{on } -I^k \\ v^k & \text{on } (-I^k)^c \end{cases}$
  - 5:  $(v^{k+1}, I^{k+1}, \delta^{k+1}) = \text{SOLVEIMPULSECONTROL}(v^{k+1/2}, (-I^k)^c)$
  - 6: **if**  $\|(v^{k+1} - v^k) / \max\{|v^{k+1}|, scale\}\|_\infty < tol$  **then**
  - 7:     break from the iteration
  - 8: **end if**
  - 9: **end for**
-

The *scale* parameter in line 5 of Algorithm 2.2.1, used throughout the literature by Forsyth, Labahn and coauthors [4,25–27], prevents the enforcement of unrealistic levels of accuracy for points  $x$  where  $v^{k+1}(x) \approx 0$ . Additionally, note that having chosen the initial guess for the payoff  $v^0$ , the initial guess for the strategy is induced by  $v^0$ . (The alternative expression for the intervention region gives the same as  $\{Mv^0 - v^0 = 0\}$  for a solution of (2.1.1).)

Line 5 of Algorithm 2.2.1 assumes we have a subroutine SOLVEIMPULSECONTROL ( $w, D$ ) that solves the constrained QVI problem:

$$\max\{Lv + f, Mv - v\} = 0 \text{ on } D, \quad \text{subject to } v = w \text{ on } D^c \tag{2.2.2}$$

for fixed  $G_{\leq 0} \subseteq D \subseteq G$  (approximate continuation region of the opponent) and  $w \in \mathbb{R}^G$  (approximate payoff due to the opponent’s intervention). Although we only need to solve for  $\tilde{v} = v_D$ , the value of  $v_{D^c} = w_{D^c}$  impacts the solution both when restricting the equations and when applying the nonlocal operator  $M$ . Hence, the approximate payoff  $v^{k+1/2}$  fed to the subroutine serves to pass on the gain that resulted from the opponent’s intervention and as an initial guess if desired (more on this on Remark 2.3.3).

The remaining of this section consists in establishing an equivalence between problem (2.2.2) and a classical (unconstrained) QVI problem of impulse control. This allows us to prove the existence and uniqueness of its solution. In particular, we will see that SOLVEIMPULSECONTROL can be defined, if wanted, by policy iteration (see the next section for an alternative method and remarks on other possible ones). Let us suppose from here onwards that the following assumptions hold true (see Appendix A for the relevant Definitions A.1.1):

- (A0) For each strategy  $\varphi = (I, \delta) \in \Phi$  and  $x \in I$ , there exists a walk in  $\text{graph}B(\delta)$  from row  $x$  to some row  $y \in I^c$ .
- (A1)  $-L$  is a strictly diagonally dominant (SDD)  $L_0$ -matrix and, for each  $\delta \in Z$ ,  $\text{Id} - B(\delta)$  is a weakly diagonally dominant (WDD)  $L_0$ -matrix.

**Remark 2.2.2 (Interpretation)** Assumptions (A0), (A1) are (H2),(H3) in [4]. For an impulse operator (say, “ $B(\delta)v(x) = v(x + \delta)$ ”), (A0) asserts that the player always wants to shift states in her intervention region to her continuation region through finitely many impulses. (This does not take into account the opponent’s response.) On the other hand, if problem (2.2.2) was rewritten as a fixed point problem, (A1) would essentially mean that the uncontrolled operator is contractive while the controlled ones are nonexpansive (see [15] and [4, Sect.4]).

**Theorem 2.2.3** *Assume (A0), (A1). Then, for every  $G_{\leq 0} \subseteq D \subseteq G$  and  $w \in \mathbb{R}^G$ , there exists a unique  $v^* \in \mathbb{R}^G$  that solves the constrained QVI problem (2.2.2). Further,  $v_D^*$  is the unique solution of*

$$\max \left\{ \tilde{L}\tilde{v} + \tilde{f}, \tilde{M}\tilde{v} - \tilde{v} \right\} = 0, \tag{2.2.3}$$

where  $\tilde{L} := L_{DD}$ ,  $\tilde{f} := f_D + L_{DD^c}w_{D^c}$ ,  $\tilde{Z} := \prod_{x \in D} Z(x)$ ,  $\tilde{c}(\tilde{\delta}) := c(\tilde{\delta}) - B(\tilde{\delta})_{DD^c}w_{D^c}$  and  $\tilde{B}(\tilde{\delta}) := B(\tilde{\delta})_{DD}$  for  $\tilde{\delta} \in \tilde{Z}$ ; and  $\tilde{M}\tilde{v} := \max_{\tilde{\delta} \in \tilde{Z}} \{ \tilde{B}(\tilde{\delta})\tilde{v} - \tilde{c}(\tilde{\delta}) \}$  for  $\tilde{v} \in \mathbb{R}^D$ .

Additionally, for any initial guess, the sequence  $(\tilde{v}^k) \subseteq \mathbb{R}^D$  defined by policy iteration [4, Thm.4.3] applied to problem (2.2.3), converges exactly to  $v_D^*$  in at most  $|\tilde{\Phi}|$  iterations, with  $\tilde{\Phi} := \{\tilde{\varphi} = (I, \tilde{\delta}) : I \subseteq G_{<0} \text{ and } \tilde{\delta} \in \tilde{Z}\}$  the set of restricted admissible strategies.<sup>18</sup>

**Proof** The equivalence between problems (2.2.2) and (2.2.3) is due to simple algebraic manipulation and  $B(\delta), c(\delta)$  being row-decoupled for every  $\delta \in Z$ .  $B(\tilde{\delta}), c(\tilde{\delta})$  are defined in the obvious way for each  $\tilde{\delta} \in \tilde{Z}$ .

The rest of the proof is mostly as in [4, Thm.4.3]. Let  $\tilde{\text{Id}} = \text{Id}_{DD}$ . Each intervention region  $I$  can be identified with its indicator  $\tilde{\psi} = \mathbb{1}_I \in \{0, 1\}^D$  since  $D \supseteq I$ . In turn, each  $\tilde{\psi}$  can be identified with the diagonal matrix  $\tilde{\Psi} = \text{diag}(\tilde{\psi}) \in \mathbb{R}^{D \times D}$ . Then, problem (2.2.3) takes the form of the classical Bellman problem

$$\max_{\tilde{\varphi} \in \tilde{\Phi}} \{ -A(\tilde{\varphi})v + b(\tilde{\varphi}) \} = 0, \tag{2.2.4}$$

if we take

$$A(\tilde{\varphi}) = -(\tilde{\text{Id}} - \tilde{\Psi})\tilde{L} + \tilde{\Psi}(\tilde{\text{Id}} - \tilde{B}(\tilde{\delta})) \quad \text{and} \quad b(\tilde{\varphi}) = (\tilde{\text{Id}} - \tilde{\Psi})\tilde{f} - \tilde{\Psi}\tilde{c}.$$

Note that  $\tilde{\Phi}$  can be identified with the Cartesian product

$$\tilde{\Phi} = \left( \prod_{x \in G_{<0}} \{0, 1\} \times Z(x) \right) \times \left( \prod_{x \in D \setminus G_{<0}} \{0\} \times Z(x) \right)$$

and  $A(\tilde{\varphi}), b(\tilde{\varphi})$  are row-decoupled for every  $\tilde{\varphi} \in \tilde{\Phi}$ . Since  $\tilde{\Phi}$  is finite, all we need to show is that the matrices  $A(\tilde{\varphi})$  are monotone (Definitions A.1.1 and [11, Thm.2.1]). Let us check the stronger property (Theorem A.1.4 and Proposition A.1.3) of being weakly chained diagonally dominant (WCDD)  $L_0$ -matrices (see Definitions A.1.1).

If (A0) and (A1) also held true for the restricted matrices and strategies, the conclusion would follow. While (A1) is clearly inherited, (A0) may fail to do so, but only in non-problematic cases. To see this, let  $\tilde{\varphi} = (I, \tilde{\delta}) \in \tilde{\Phi}$ ,  $x \in I \subseteq D$  and let  $\delta \in Z$  be some extension of  $\tilde{\delta}$ . Note that row  $x$  of  $A(\tilde{\varphi})$  is WDD. We want to show that there is a walk in  $\text{graph}A(\tilde{\varphi})$  from  $x$  to an SDD row.

By (A0) there must exist some walk  $x = y_0 \rightarrow y_1 \rightarrow \dots \rightarrow y_n \in I^c$  in  $\text{graph}B(\delta)$ . If this is in fact a walk from  $x$  to  $I^c \cap D$  in  $\text{graph}\tilde{B}(\tilde{\delta})$ , then it verifies the desired property (just as in [4, Thm.4.3]). If not, then there must be a first  $0 \leq m < n$  such that the subwalk  $x \rightarrow \dots \rightarrow y_m$  is in  $\text{graph}\tilde{B}(\tilde{\delta})$  but  $y_{m+1} \notin D$ . Since  $y_m \rightarrow y_{m+1}$  is an edge in  $\text{graph}B(\delta)$ , we have  $B(\delta)_{y_m, y_{m+1}} \neq 0$  and the WDD row (by (A1))  $y_m$  of  $\tilde{\text{Id}} - \tilde{B}(\tilde{\delta})$  is in fact SDD. Meaning that the subwalk  $x \rightarrow \dots \rightarrow y_m$  verifies the desired property instead. □

**Remark 2.2.4 (Practical considerations)** 1. While convergence is guaranteed to be exact, floating point arithmetic can bring about stagnation [27]. A stopping criteria like that of Algorithm 2.2.1 should be used in those cases, with a tolerance  $\ll tol$ . 2. The

<sup>18</sup>  $|A|$  denotes the cardinal of set  $A$ .

solution of system (2.2.3) does not change if one introduces a scaling factor  $\lambda > 0$ :  $\max \{ \tilde{L}\tilde{v} + \tilde{f}, \lambda(\tilde{M}\tilde{v} - \tilde{v}) \} = 0$  [4, Lem.4.1]. This problem-specific parameter is typically added in the implementation to enhance performance [4,26]. It can intuitively be thought as a units adjustment.

### 2.3 Iterative Subroutine for Impulse Control

Due to Theorem 2.2.3, a sensible choice for SOLVEIMPULSECONTROL is the classical policy iteration algorithm [4, Thm.4.3] applied to (2.2.3) (i.e., [11, Ho-1] applied to (2.2.4)), adding an appropriately chosen scaling factor  $\lambda$  to improve efficiency (Remark 2.2.4 2). One possible drawback of such choice is the following: at each iteration, one needs to solve the system  $-A(\tilde{\varphi}^k)v^{k+1} + b(\tilde{\varphi}^k) = 0$  for some  $\tilde{\varphi}^k \in \tilde{\Phi}$ . While the matrix  $\tilde{L}$  typically has a good sparsity pattern in applications (often tridiagonal), the presence of  $\tilde{B}(\tilde{\delta}^k)$  prevents  $A(\tilde{\varphi}^k)$  from inheriting the same structure in general, and makes the resolution of the previous system more costly. (See [26] where this issue is addressed for HJB problems with jump diffusions and regime switching, among others.)

Motivated by the previous observation, this section considers an alternative choice for SOLVEIMPULSECONTROL: an instance of a very general class of algorithms known as *fixed-point policy iteration* [14,26]. As far as the author knows, this application to impulse control was never done in the past other than heuristically in [3]. Instead of solving  $-A(\tilde{\varphi}^k)v^{k+1} + b(\tilde{\varphi}^k) = 0$  at the  $k$ -th iteration, we will solve

$$\underbrace{(\tilde{\text{Id}} - \tilde{\Psi}^k)\tilde{L} - \tilde{\Psi}^k}_{-\tilde{\mathbb{A}}(\tilde{\varphi}^k)} v^{k+1} + \underbrace{\tilde{\Psi}^k \tilde{B}(\tilde{\delta}^k)}_{\tilde{\mathbb{B}}(\tilde{\varphi}^k)} v^k + \underbrace{b(\tilde{\varphi}^k)}_{\tilde{\mathbb{C}}(\tilde{\varphi}^k)} = 0, \tag{2.3.1}$$

(scaled by  $\lambda$ ) where the previous iterate value  $v^k$  is given and  $\tilde{\Psi}^k$  is the diagonal matrix with  $\psi^k$  as diagonal. In other words, we split the original policy matrix  $A(\tilde{\varphi}) = \tilde{\mathbb{A}}(\tilde{\varphi}) - \tilde{\mathbb{B}}(\tilde{\varphi})$  and we apply a one-step fixed-point approximation,

$$\tilde{\mathbb{A}}(\tilde{\varphi}^k)\tilde{v}^{k+1} = \tilde{\mathbb{B}}(\tilde{\varphi}^k)\tilde{v}^k + \tilde{\mathbb{C}}(\tilde{\varphi}^k), \tag{2.3.2}$$

at each iteration of Howard’s algorithm. The resulting method can be expressed as follows (*tol* and *scale* are as in Algorithm 2.2.1):

**Subroutine 2.3.1**  $(v, I, \delta) = \text{SOLVEIMPULSECONTROL}(w, D)$

**Inputs**  $w \in \mathbb{R}^G$  and solvency region  $G_{\leq 0} \subseteq D \subseteq G$

**Outputs**  $v \in \mathbb{R}^G, I \subseteq G_{<0}, \delta \in Z$

Set scaling factor  $\lambda > 0$  and  $0 < \tilde{tol} \ll tol$ .

// Restrict constrained problem

1:  $\tilde{L} := L_{DD}, \tilde{f} := f_D + L_{DD^c} w_{D^c}$

2:  $\tilde{Z} := \prod_{x \in D} Z(x), \tilde{c}(\delta) := c(\delta) - B(\delta)_{DD^c} w_{D^c}, \tilde{B}(\delta) := B(\delta)_{DD}$  for  $\delta \in \tilde{Z}$

3:  $\tilde{M}\tilde{v} := \max_{\delta \in \tilde{Z}} \{ \tilde{B}(\delta)\tilde{v} - \tilde{c}(\delta) \}$  for  $\tilde{v} \in \mathbb{R}^D, \text{Id} := \text{Id}_{DD}$

// Solve restricted problem

4: Choose initial guess:  $\tilde{v}^0 \in \mathbb{R}^D, I^0 \subseteq G_{<0}$

5: **for**  $k = 0, 1, 2, \dots$  **do**

6:  $\tilde{L}^k_{ij} = \begin{cases} \tilde{L}_{ij} & \text{if } x_i \in D \setminus I^k \\ -\text{Id}_{ij} & \text{if } x_i \in I^k \end{cases} \quad \tilde{f}^k = \begin{cases} \tilde{f} & \text{on } D \setminus I^k \\ \tilde{M}\tilde{v}^k & \text{on } I^k \end{cases}$

7:  $\tilde{v}^{k+1}$  solution of  $\tilde{L}^k \tilde{v} + \tilde{f}^k = 0$

8:  $I^{k+1} = \{ \tilde{L}\tilde{v}^{k+1} + \tilde{f} \leq \lambda(\tilde{M}\tilde{v}^{k+1} - \tilde{v}^{k+1}) \}$

9: **if**  $\|(\tilde{v}^{k+1} - \tilde{v}^k) / \max\{\tilde{v}^{k+1}, \text{scale}\}\|_\infty < \tilde{tol}$  **then**

10:  $v = \begin{cases} \tilde{v}^{k+1} & \text{on } D \\ w_{D^c} & \text{on } D^c \end{cases}, I = I^{k+1}, \delta = \delta^*(v)$  and break from the iteration

11: **end if**

12: **end for**

Lines 1–3 of Sect. 2.3.1 deal with restricting the constrained problem, while the rest give a routine that can be applied to any QVI of the form (2.2.3). Starting from some suboptimal  $\tilde{v}^0$  and  $I^0$ , one computes a new payoff  $\tilde{v}^1$  by solving the coupled equations  $\tilde{M}\tilde{v}^0 - \tilde{v}^1 = 0$  on  $I^0$  and  $\tilde{L}\tilde{v}^1 + \tilde{f} = 0$  outside  $I^0$ . A new intervention region  $I^1 = \{ \tilde{L}\tilde{v}^1 + \tilde{f} \leq \lambda(\tilde{M}\tilde{v}^1 - \tilde{v}^1) \}$  is defined and the procedure is iterated.

Algorithmically, the difference with classical policy iteration is that  $\tilde{v}^{k+1}$  is computed in Line 7 with a fixed obstacle  $\tilde{M}\tilde{v}^k$ , changing a quasivariational inequality for a variational one. The resulting method is intuitive and simple to implement, and the linear system (2.3.1) (Line 7) inherits the sparsity pattern of  $\tilde{L}$ . For example, for an SDD tridiagonal  $\tilde{L}$ , the system can be solved (exactly in exact arithmetic or stably in floating point one) in  $O(n)$  operations, with  $n = |D|$  [28, Sect.9.5]. The matrix-vector multiply  $\tilde{B}(\delta^k)\tilde{v}^k$  can take at most  $O(n^2)$  operations, but will reduce to  $O(n)$  for standard discretizations of impulse operators.

It is also worth mentioning that Sect. 2.3.1 differs from the so-called iterated optimal stopping [16,33] in that the latter solves  $\max \{ \tilde{L}\tilde{v}^{k+1} + \tilde{f}, \tilde{M}\tilde{v}^k - \tilde{v}^{k+1} \} = 0$  exactly at the  $k$ -th iteration (by running a full subroutine of Howard’s algorithm with fixed obstacle), while the former only performs one approximation step.

To establish the convergence of Sect. 2.3.1 in the present framework, we add the following assumption:

(A2)  $B(\delta)$  has nonnegative diagonal elements for all  $\delta \in Z$ .

**Remark 2.3.1** (A2) and the requirement of (A1) that  $\text{Id} - B(\delta)$  be a WDD  $L_0$ -matrix are equivalent to  $B(\delta)$  being substochastic (see Appendix A). This is standard for

impulse operators (see Sect. 2.6) and other applications of fixed-point policy iteration [26, Sect.4-5].

**Theorem 2.3.2** *Assume (A0)–(A2) and set  $I^0 = \emptyset$ . Then, for every  $G_{\leq 0} \subseteq D \subseteq G$  and  $w \in \mathbb{R}^G$ , the sequence  $(\tilde{v}^k)$  defined by SOLVEIMPULSECONTROL( $w, D$ ) is monotone increasing for  $k \geq 1$  and converges to the unique solution of (2.2.3).*

**Proof** We can assume without loss of generality that  $\lambda = 1$ . Sect. 2.3.1 takes the form of a fixed-point policy iteration algorithm as per (2.3.2). Assumptions (A0), (A1) ensure (2.2.3) has a unique solution (Theorem 2.3.1) and that this scheme satisfies [26, Cond.3.1 (i),(ii)]. That is,  $\tilde{\mathbb{A}}(\tilde{\varphi})$  and  $\tilde{\mathbb{A}}(\tilde{\varphi}) - \tilde{\mathbb{B}}(\tilde{\varphi})$  are nonsingular  $M$ -matrices (see proof of Theorem 2.3.1 and Appendix A) and all coefficients are bounded since  $\tilde{\Phi}$  is finite. In [26, Thm.3.4] convergence is proved under one additional assumption of  $\|\cdot\|_\infty$ -contractiveness [26, Cond.3.1 (iii)], which is not verified in our case. However, the same computations show that the scheme satisfies

$$\tilde{\mathbb{A}}(\tilde{\varphi}^k)(\tilde{v}^{k+1} - \tilde{v}^k) \geq \tilde{\mathbb{B}}(\tilde{\varphi}^{k-1})(\tilde{v}^k - \tilde{v}^{k-1}) \text{ for all } k \geq 1. \tag{2.3.3}$$

Since  $I^0 = \emptyset$ , and due to (A1) and (A2),  $\tilde{\mathbb{B}}(\tilde{\varphi}^0) = 0$  and  $\tilde{\mathbb{B}}(\tilde{\varphi}^k) \geq 0$  for all  $k$ . Thus,  $(\tilde{v}^k)_{k \geq 1}$  is increasing by monotonicity of  $\tilde{\mathbb{A}}(\tilde{\varphi}^k)$ . Furthermore, it must be bounded, since for all  $k \geq 1$ :

$$\tilde{\mathbb{A}}(\tilde{\varphi}^k)\tilde{v}^{k+1} = \tilde{\mathbb{B}}(\tilde{\varphi}^k)\tilde{v}^k + \tilde{\mathbb{C}}(\tilde{\varphi}^k) \leq \tilde{\mathbb{B}}(\tilde{\varphi}^k)\tilde{v}^{k+1} + \tilde{\mathbb{C}}(\tilde{\varphi}^k),$$

which gives  $\tilde{v}^{k+1} \leq (\tilde{\mathbb{A}}(\tilde{\varphi}^k) - \tilde{\mathbb{B}}(\tilde{\varphi}^k))^{-1}\tilde{\mathbb{C}}(\tilde{\varphi}^k) \leq \max_{\tilde{\varphi} \in \tilde{\Phi}} (\tilde{\mathbb{A}}(\tilde{\varphi}) - \tilde{\mathbb{B}}(\tilde{\varphi}))^{-1}\tilde{\mathbb{C}}(\tilde{\varphi})$ . Hence,  $(\tilde{v}^k)$  converges. That the limit solves (2.2.3) is proved as in [26, Lem.3.3].  $\square$

**Remark 2.3.3** In light of Theorem 2.3.2, moving forward we will set  $I^0 = \emptyset$  in Sect. 2.3.1. (Note that the value of  $\tilde{v}^0$  is irrelevant in this case, since  $\tilde{\mathbb{B}}(\tilde{\varphi}^0) = 0$ .) It is natural however to choose  $\tilde{v}^0 = w_D$  and  $I^0 = \{\tilde{L}\tilde{v}^0 + \tilde{f} \leq \lambda(\tilde{M}\tilde{v}^0 - \tilde{v}^0)\}$ . The experiments performed with the latter choice displayed faster but non-monotone convergence, but this is not proved here. Additionally, exact convergence was often observed.

**Remark 2.3.4** For the experiments in Sect. 3, SOLVEIMPULSECONTROL was chosen as Sect. 2.3.1 instead of classical policy iteration, since the former displayed overall lower runtimes when compared to the latter. However, it should be noted that the games considered have relatively large costs, with many parameter values taken from previous works [1,3,10]. For smaller costs often occurring in practice, and especially for large-scale problems, the convergence rate of Sect. 2.3.1 can become very slow, as it happens with iterated optimal stopping [34, Rmk.3.1]. This is not so for policy iteration, which converges superlinearly (see [11, Thm.3.4] and [34, Sect.5]), making it a better suited choice in this case (in terms of runtime as well). Furthermore, it has been demonstrated in [35][Sect.7] that the number of steps required for penalized policy iteration to converge remains bounded as the grid is refined. While penalized methods are not considered in the present work, such mesh-independence properties are desirable and could prove paramount in further studies of Algorithm 2.2.1 (see end of Sect. 2.5 and Sect. 3).

### 2.4 Overall Routine as a Fixed-Point Policy-Iteration-Type Method

The system of QVIs (2.1.1) cannot be reduced in any apparent way to a Bellman formulation (2.1.4) (see comments preceding equation). Notwithstanding, we shall see that Algorithm 2.2.1 does take a very similar form to a fixed-point policy iteration algorithm as in (2.3.2) for some appropriate  $\mathbb{A}, \mathbb{B}, \mathbb{C}$ . Further, assumptions resembling those of the classical case [26] will be either satisfied or imposed to study its convergence. The matrix and graph-theoretic definitions and properties used throughout this section can be found in Appendix A.

**Notation 2.4.1** We identify each intervention region  $I \subseteq G_{<0}$  with its indicator function  $\psi = \mathbb{1}_I \in \{0, 1\}^G$  and each  $\psi$  with the diagonal matrix  $\Psi = \text{diag}(\psi) \in \mathbb{R}^{G \times G}$ . The sequences  $(v^k)$  and  $(\varphi^k)$ , with  $\varphi^k = (\psi^k, \delta^k)$ , are the ones generated by Algorithm 2.2.1. SOLVEIMPULSECONTROL is defined as either Sect. 2.3.1 or Howard’s algorithm (Theorem 2.2.3), setting the outputs  $I, \delta$  as done for the former subroutine. We consider  $v^* \in \mathbb{R}^G$  fixed and  $\varphi^* = (\psi^*, \delta^*(v^*))$  the induced strategy with  $\psi^* := \{Lv^* + f \leq Mv^* - v^*\} \cap G_{<0}$ .

**Proposition 2.4.2** Assume (A0)–(A2). Then,

$$\mathbb{A}(\varphi^k, \varphi^{k+1})v^{k+1} = \mathbb{B}(\varphi^k)v^k + \mathbb{C}(\varphi^k, \varphi^{k+1}), \text{ where:} \tag{2.4.1}$$

- (i)  $\psi^k = \mathbb{1}_{\{Lv^k + f \leq Mv^k - v^k\} \cap G_{<0}}$  and  $\delta^k \in \arg \max_{\delta \in Z} \{B(\delta)v^k - c(\delta)\}$ .
- (ii)  $\mathbb{A}(\varphi, \bar{\varphi}) := \text{Id} - (\text{Id} - \bar{\Psi} - S\Psi S)(\text{Id} + L) - \bar{\Psi}B(\bar{\delta})$  is a WCDD  $L_0$ -matrix, and thus a nonsingular  $M$ -matrix.
- (iii)  $\mathbb{B}(\varphi) := S\Psi B(\delta)S = \text{diag}(S\psi)SB(\delta)S$  is substochastic.
- (iv)  $\mathbb{C}(\varphi, \bar{\varphi}) := (\text{Id} - \bar{\Psi} - S\Psi S)f - \bar{\Psi}c(\bar{\delta}) + S\Psi Sg(S\delta)$ .

**Proof** Using that  $(v^{k+1}, I^{k+1}, \delta^{k+1}) = \text{SOLVEIMPULSECONTROL}(v^{k+1/2}, (-I^k)^c)$  solves the constrained QVI problem (2.2.2) for  $D = (-I^k)^c$  and  $w = v^{k+1/2}$  (Theorems 2.2.3 or 2.3.2), the recurrence relation (2.4.1) results from simple algebraic manipulation.

Given  $\varphi, \bar{\varphi} \in \Phi$ , (A0) and (A1) ensure  $\mathbb{A}(\varphi, \bar{\varphi})$  is a WCDD  $L_0$ -matrix, while (A1) and (A2) imply  $\mathbb{B}(\varphi)$  is substochastic.

The following corollary is immediate by induction. It gives a representation of the sequence of payoffs in terms of the improving strategies throughout the algorithm.

**Corollary 2.4.3** Assume (A0)–(A2). Then,<sup>19</sup>

$$v^{k+1} = \left( \prod_{j=k}^0 \mathbb{A}^{-1} \mathbb{B}(\varphi^j, \varphi^{j+1}) \right) v^0 + \sum_{n=0}^k \left( \prod_{j=k}^{n+1} \mathbb{A}^{-1} \mathbb{B}(\varphi^j, \varphi^{j+1}) \right) \mathbb{A}^{-1} \mathbb{C}(\varphi^n, \varphi^{n+1}). \tag{2.4.2}$$

<sup>19</sup> For any index  $i \leq k, \prod_{j=k}^i A^j := A^k A^{k-1} \dots A^i$ .

We now establish some properties of the strategy-dependent matrix coefficients that will be useful in the sequel. Given a WDD (resp. substochastic) matrix  $A \in \mathbb{R}^{G \times G}$ , we define its set of “non-trouble states” (or rows) as

$$J[A] := \{x \in G : \text{row } x \text{ of } A \text{ is SDD}\} \text{ (resp. } \hat{J}[A] := \{x \in G : \text{row } x \text{ of } A \text{ sums less than one}\},$$

and its *index of connectivity*  $\text{con}A$  (resp. *index of contraction*  $\widehat{\text{con}}A$ ) by computing for each state the least length that needs to be walked on  $\text{graph}A$  to reach a non-trouble one, and then taking the maximum over all states (more details in Appendix A). This recently introduced concept gives an equivalent characterization of the WCDD property for a WDD matrix as  $\text{con}A < +\infty$ , and can be efficiently checked for sparse matrices in  $O(|G|)$  operations [7]. On the other hand, if  $A$  is substochastic then  $\widehat{\text{con}}A < \infty$  if and only if its spectral radius verifies  $\rho(A) < 1$  (Theorem A.1.6). The proof of the following lemma can be found in Appendix A.

**Lemma 2.4.4** *Assume (A0)–(A2). Then for all  $\varphi, \bar{\varphi} \in \Phi$ ,  $\mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})$  is substochastic,  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is a WDD  $L_0$ -matrix and  $\widehat{\text{con}}[\mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})] \leq \text{con}[(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})]$ .*

As previously mentioned, system (2.1.1) may have no solution. The matrix coefficients introduced in this section allow us to algebraically characterize the existence of such solutions through strategy-dependent linear systems of equations.

**Proposition 2.4.5** *Assume (A0)–(A2). Then the following statements are equivalent:*

- (i)  $v^*$  solves the system of QVIs (2.1.1).
- (ii)  $\mathbb{A}(\varphi^*, \varphi^*)v^* = \mathbb{B}(\varphi^*)v^* + \mathbb{C}(\varphi^*, \varphi^*)$ .

As mentioned in Remark 2.2.2, Assumption (A0) constrains the type of strategies the player can use, but without taking into account the opponent’s response. This is enough for the single-player constrained problems to have a solution and, therefore, for Algorithm 2.2.1 to be well defined. But we cannot expect this restriction to be sufficient in the study of the two-player game and the convergence of the overall routine.

In order to improve the result of Proposition 2.4.5 let us consider the following stronger version of (A0) reflecting the interaction between the player and the opponent.

(A1’) For each pair of strategies  $\varphi, \bar{\varphi} \in \Phi$ , and for each  $x \in \bar{I} \cup (-I)$ , there exists a walk in  $\text{graph}(\bar{\Psi}B(\bar{\delta}) + S\Psi B(\delta)S)$  from row  $x$  to some row  $y \in \bar{C} \cap C$ , where  $\bar{C} = \bar{I}^c$ ,  $C = I^c$ .

**Remark 2.4.6 (Interpretation)** If  $\bar{\varphi}, -\varphi$  are the strategies used by the player and the opponent respectively,<sup>20</sup> then (A1’) asserts that states in their intervention regions will eventually be shifted to the common continuation region. This precludes infinite simultaneous interventions and emulates the admissibility condition of the continuous-state case. Fixing  $I = \emptyset$  we recover (A0). Additionally, (A1’) together with (A1) imply

<sup>20</sup> The slight abuse of notation  $-\varphi$  stands for the strategy symmetric to  $\varphi$ , i.e.,  $-\varphi = (-I, -\delta(-x))$ .

that  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is a WCDD  $L_0$ -matrix,<sup>21</sup> hence an  $M$ -matrix. This is another one of the assumptions of the classical fixed-point policy iteration [26].

Under this new assumption, the  $\varphi^* = \varphi^*(v^*)$ -dependent systems of Proposition 2.4.5 will admit a unique solution. Then solving the original problem (2.1.1) amounts to finding  $v^* \in \mathbb{R}^G$  that solves its induced linear system of equations.

**Proposition 2.4.7** *Assume (A1'), (A1), (A2). In the context of Proposition 2.4.5, the following statements are also equivalent:*

- (iv)  $v^* = (\mathbb{A} - \mathbb{B})^{-1} \mathbb{C}(\varphi^*, \varphi^*)$ .
- (v)  $v^* = (\text{Id} - \mathbb{A}^{-1} \mathbb{B})^{-1} \mathbb{A}^{-1} \mathbb{C}(\varphi^*, \varphi^*) = \sum_{n \geq 0} (\mathbb{A}^{-1} \mathbb{B})^n \mathbb{A}^{-1} \mathbb{C}(\varphi^*, \varphi^*)$ . (cf. equation (2.4.2).)

**Proof** Both expressions result from rewriting and solving the systems of Proposition 2.4.5. Assumptions (A1'), (A1) guarantee that  $(\mathbb{A} - \mathbb{B})(\varphi^*, \varphi^*)$  is WCDD (Remark 2.4.6) and, in particular, nonsingular. Then (v) is due to Lemma 2.4.4, Theorem A.1.6 and the matrix power series expansion  $(\text{Id} - X)^{-1} = \sum_{n \geq 0} X^n$ , when  $\rho(X) < 1$ . □

### 2.5 Convergence Analysis

We now study the convergence properties of Algorithm 2.2.1. Henceforth, the UIP refers to the obvious discrete analogous of Definition 1.2.1, where we replace the domain  $\mathbb{R}$ , the impulse constraints  $\mathcal{Z}$  and the operator  $\mathcal{M}$  by their discretizations  $G$ ,  $Z$  and  $M$  respectively.

The obvious first question to address is whether when Algorithm 2.2.1 converges, it does so to a solution of the system of QVIs (2.1.1). Unlike in the classical Bellman problem (2.1.4), problem (2.1.1) is intrinsically dependent on the particular strategy chosen by the player (see Propositions 2.4.5 and 2.4.7). Accordingly, we start with a lemma addressing what can be said about the convergence of the strategies  $(\varphi^k)$  when the payoffs  $(v^k)$  converge.

**Notation 2.5.1**  $\partial I^* := \{Lv^* + f = Mv^* - v^*\} \cap G_{<0}$  denotes the “border” of the intervention region  $\{Lv^* + f \leq Mv^* - v^*\} \cap G_{<0}$  defined by  $v^*$ .

**Lemma 2.5.2** *Assume (A0)–(A2) and suppose  $v^k \rightarrow v^*$ . Then:*

- (i)  $\psi^k \rightarrow \psi^*$  in  $(\partial I^*)^c$  and  $Mv^k \rightarrow Mv^*$ .
- (ii) If  $\bar{\psi}, \bar{\delta}$  are any two limit points of  $(\psi^k), (\delta^k)$  resp.,<sup>22</sup> then

$$\bar{\delta} \in \arg \max_{\delta \in Z} \{B(\delta)v^* - c(\delta)\}, \quad \bar{\psi} = 0 \text{ on } G_{>=0} \quad \text{and} \quad \bar{\psi} \in \arg \max_{i \in \{0,1\}} \{O_i v^*\} \text{ on } G_{<0},$$

<sup>21</sup> (A1) guarantees the WDD  $L_0$ -property (see proof of Lemma 2.4.4) with SDD  $i$ -th row for any  $x_i \in \bar{C} \cap C$ , as it coincides with the  $i$ -th row of  $-L$ . Since  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  and  $\text{Id} - (\bar{\Psi}B(\bar{\delta}) + S\Psi B(\delta)S)$  agree for any row indexed in  $\bar{I} \cup (-I)$  and  $\text{graph}(\text{Id} - (\bar{\Psi}B(\bar{\delta}) + S\Psi B(\delta)S))$  is equal to  $\text{graph}(\bar{\Psi}B(\bar{\delta}) + S\Psi B(\delta)S)$  except possibly for self loops, adding (A1') means that any WDD row of  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is connected to some SDD row.

<sup>22</sup> By “limit point” we mean the limit of a convergent subsequence.

with  $O_0v = Lv + f$  and  $O_1v = Mv - v$ .  
 (iii) If  $v^*$  has the UIP, then  $\delta^k \rightarrow \delta^*(v^*)$  and  $Hv^k \rightarrow Hv^*$ .

**Proof** That  $Mv^k \rightarrow Mv^*$  is clear by continuity of the operators  $B(\delta)$  and finiteness of  $Z$ .

Let  $x \in (\partial I^*)^c$  and suppose  $Lv^*(x) + f(x) < Mv^*(x) - v^*(x)$  (the other case being analogous). By continuity of  $L$  and  $M$  there must exist some  $k_0$  such that  $Lv^k(x) + f(x) < Mv^k(x) - v^k(x)$  for all  $k \geq k_0$ , which implies  $\psi^k(x) = 1 = \psi^*(x)$  for  $k \geq k_0$ .

The statement about  $\bar{\psi}, \bar{\delta}$  is proved as before by considering appropriate subsequences. Consequently, if  $v^*$  has the UIP, then necessarily  $\delta^k \rightarrow \delta^*(v^*)$  and  $Hv^k \rightarrow Hv^*$ . □

As a corollary we can establish that, should the sequence  $(v^k)$  converge, its limit must solve problem (2.1.1). If convergence is not exact however (i.e., in finite iterations), then we will ask that  $v^*$  verifies some of the properties of the Verification Theorem in Corollary 1.3.9. Namely, the UIP and a discrete analogous of the continuity in the border of the opponent’s intervention region. We emphasise that our main motivation in solving system (2.1.1) relies in Corollary 1.3.9 and its framework. Additionally, in most practical situations and for fine-enough grids, one can intuitively expect the discretization of an equilibrium payoff as in Corollary 1.3.9 to inherit the UIP. Lastly, we note that the exact equality  $Lv^* + f = Mv^* - v^*$  will typically not be verified for any point in the grid in practice, giving  $\partial I^* = \emptyset$ .

**Corollary 2.5.3** Assume (A0)– (A2) and suppose  $v^k \rightarrow v^*$ . Then:

- (i) If the convergence is exact, then  $v^*$  solves the system of QVIs (2.1.1).
- (ii) If  $v^*$  has the UIP and  $Lv^* + f = Hv^* - v^*$  on  $-\partial I^*$ , then  $v^*$  solves (2.1.1).

**Proof** (i) is immediate from the definition of Algorithm 2.2.1.

In the general case, since  $\{0, 1\}^G$  is finite, there is a subsequence of  $(\psi^k, \psi^{k+1})$  that converges to some pair  $(\psi, \bar{\psi})$ . Passing to such subsequence, by Lemma 2.5.2, the UIP of  $v^*$  and equation (2.4.1), we get that  $v^*$  solves the system  $\mathbb{A}(\varphi, \bar{\varphi})v^* = \mathbb{B}(\varphi)v^* + \mathbb{C}(\varphi, \bar{\varphi})$  for  $\varphi = (\psi, \delta^*(v^*))$ ,  $\bar{\varphi} = (\bar{\psi}, \delta^*(v^*))$  and  $\psi, \bar{\psi}$  coincide with  $\psi^*$  except possibly on  $\partial I^*$ . Thus, it only remains to show that  $v^*$  also solves the equations of the system (2.1.1) for any  $x \in \partial I^* \cup (-\partial I^*)$ .

For  $x \in \partial I^*$ , the previous is true by definition. Suppose now  $x \in -\partial I^* \subseteq \bar{I}^c$ . We have  $\psi^*(-x) = 1$ . If  $\psi(-x) = 1$ , there is nothing to prove. If  $\psi(-x) = 0$ , then  $x \in \bar{I}^c \cap (-I)^c$  and  $0 = Lv^*(x) + f(x) = Hv^*(x) - v^*(x)$ , where the last equality holds true by assumption. □

Lemma 2.5.2 shows to what extent the convergence of the payoffs imply the convergence of the strategies. The following theorem, of theoretical interest, establishes a reciprocal under the stronger assumption (A1’). In general, since the set of strategies  $\Phi$  is finite, the sequence of strategy-dependent coefficients of the fixed-point equations (2.4.1) will always be bounded and with finitely many limit points. However, if the approximating strategies are such that the former coefficients convergence, then

Algorithm 2.2.1 is guaranteed to converge. Further, instead of looking at the convergence of  $(\mathbb{A}, \mathbb{B}, \mathbb{C})(\varphi^k, \varphi^{k+1})$ , we can instead consider the weaker condition of  $(\mathbb{A}^{-1}\mathbb{B}, \mathbb{A}^{-1}\mathbb{C})(\varphi^k, \varphi^{k+1})$  converging.

**Theorem 2.5.4** *Assume (A1'), (A1), (A2). If  $(\mathbb{A}^{-1}\mathbb{B}(\varphi^k, \varphi^{k+1}))$  and  $(\mathbb{A}^{-1}\mathbb{C}(\varphi^k, \varphi^{k+1}))$  converge, then  $(v^k)$  converges.*

**Proof** Set  $b = \lim_k \mathbb{A}^{-1}\mathbb{C}(\varphi^k, \varphi^{k+1})$ . Since  $\Phi$  is finite, there must exist  $k_0 \in \mathbb{N}$  and  $\varphi, \bar{\varphi} \in \Phi$  such that  $\mathbb{A}^{-1}\mathbb{B}(\varphi^k, \varphi^{k+1}) = \mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})$  and  $\mathbb{A}^{-1}\mathbb{C}(\varphi^k, \varphi^{k+1}) = b$  for all  $k \geq k_0$ . Moreover, under our assumptions,  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is a WCDD  $L_0$ -matrix. Then Lemma 2.4.4 and Theorem A.1.6 imply that  $\mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})$  is contractive for some matrix norm. Lastly, note that the sequence of payoffs  $(v^k)_{k \geq k_0}$  now satisfies the classical (constant-coefficients) contractive fixed-point recurrence  $v^{k+1} = \mathbb{A}^{-1}\mathbb{B}(\varphi, \bar{\varphi})v^k + b$ , which converges to the unique fixed-point of the equation.  $\square$

The classical fixed-point policy-iteration framework [14,26] assumes uniform contractiveness in  $\|\cdot\|_\infty$  of the sequence of operators. This is a natural norm to consider in a context where matrices have properties defined row by row, such as diagonal dominance.<sup>23</sup> However, the authors mention convergence in experiments where only  $\|\cdot\|_\infty$ -nonexpansiveness held true. The latter is the typical case in our context, for the matrices  $\mathbb{A}^{-1}\mathbb{B}(\varphi^k, \varphi^{k+1})$ , which is why Theorem 2.5.4 relies on the fact that a spectral radius strictly smaller than one guarantees contractiveness in some matrix norm.

It is natural to ask whether there is some contractiveness condition that may account for the observations in [14,26] and that can be generalized to our context to further the study of Algorithm 2.2.1. Imposing a uniform bound on the spectral radii would not only be hard to check, but also difficult to manipulate, as the spectral radius is not sub-multiplicative.<sup>24</sup> Instead, we can consider the sequential indices of contraction and connectivity, which naturally generalize those of the previous section by means of walks in the graph of a sequence of matrices (see Appendix A for more details). As before, they can be identified with one another (see Lemma A.1.5) and, given substochastic matrices, the sequential index of contraction tells us how many we need to multiply before the result becomes  $\|\cdot\|_\infty$ -contractive (Theorem A.1.7). Thus, let us consider a uniform bound on the following sequential indices of connectivity:

(A0'') There exists  $m \in \mathbb{N}_0$  such that for any sequence of strategies  $(\bar{\varphi}^k) \subseteq \Phi$ ,

$$\text{con} \left[ \left( (\mathbb{A} - \mathbb{B})(\bar{\varphi}^k, \bar{\varphi}^{k+1}) \right) \right] \leq m.$$

**Remark 2.5.5** Given  $\varphi, \bar{\varphi} \in \Phi$ , by considering the sequence  $\varphi, \bar{\varphi}, \varphi, \bar{\varphi}, \dots$ , we see that (A0'') implies (A1'). In fact, (A0'') can be interpreted as precluding infinite simultaneous impulses even when the players can adapt their strategies (cf. Remark 2.4.6) and imposing that the number of shifts needed for any state to reach the common continuation region is bounded.

<sup>23</sup> Recall that this norm can be computed as the maximum absolute value row sum.

<sup>24</sup>  $\rho(AB) \leq \rho(A)\rho(B)$  does not hold in general when the matrices  $A$  and  $B$  do not commute.

Under this stronger assumption, we have:

**Proposition 2.5.6** *Assume (A0''), (A1), (A2). Then  $(v^k)$  is bounded.*

**Proof** In a similar way to Lemma 2.4.4, one can check that under (A0''), (A1), (A2) we have the following uniform bound for the sequential indices of contraction:

$$\widehat{\text{con}} \left[ \left( \mathbb{A}^{-1} \mathbb{B}(\varphi^k, \varphi^{k+1}) \right) \right] \leq m,$$

for any sequence of strategies  $(\varphi^k) \subseteq \Phi$ . In other words, not only is any product of the previous substochastic matrices also substochastic (i.e.,  $\|\cdot\|_\infty$ -nonexpansive), but it is also  $\|\cdot\|_\infty$ -contractive when there are at least  $m + 1$  factors. Furthermore, since  $\Phi^{m+1}$  is finite, there must be a uniform contraction constant  $C_1 < 1$ . Let  $C_2 > 0$  be a uniform bound for  $\mathbb{A}^{-1}C$ . By the representation (2.4.2)<sup>25</sup>

$$\|v^{k+1}\|_\infty \leq \|v^0\|_\infty + C_2 \sum_{n=0}^k C_1^{\lfloor \frac{k-n}{m+1} \rfloor} \leq \|v^0\|_\infty + (m + 1)C_2 \sum_{n=0}^\infty C_1^n < +\infty.$$

□

Given  $n_0 \in \mathbb{N}$  and  $k > (m + 1)n_0$ , the same argument of the previous proof shows that one can decompose  $(v^k)$  as  $v^{k+1} = u^k + F(\varphi^{k-(m+1)n_0}, \dots, \varphi^k) + w^k$ , for a fixed function  $F$ ,  $\|u^k\|_\infty \leq C_1^{\lfloor k/(m+1) \rfloor} \|v^0\| \rightarrow 0$  and  $\|w^k\|_\infty \leq (m + 1)C_2 \sum_{n=n_0}^\infty C_1^n$ . The latter is small if  $n_0$  is large. Hence, one could heuristically expect that the trailing strategies are often the ones dominating the convergence of the algorithm. In fact, in all the experiments carried out with a discretization satisfying (A0''), (A1), (A2), a dichotomous behaviour was observed: the algorithm either converged or at some point reached a cycle between a few payoffs. In the latter case, and restricting attention to instances in which one heuristically expects a solution to exist (more details in Sect. 3), it was possible to reduce the residual to the QVIs and the distance between the iterates by refining the grid.

The previous motivates the study of Algorithm 2.2.1 when the grid is sequentially refined, instead of fixed. Such an analysis however, would likely entail the need of a viscosity solutions framework as in [2,12], which does not currently exist in the literature of nonzero-sum stochastic impulse games. Consequently, this analysis and the stronger convergence results that may come out of it are inevitably outside the scope of this paper.

## 2.6 Discretization Schemes

Let us conclude this section by showing how one can discretize the symmetric system of QVIs (1.3.1) to obtain (2.1.1) in a way that satisfies the assumptions present throughout the paper. Recall that we work on a given symmetric grid  $G : x_{-N} = -x_N < \dots < x_{-1} = -x_1 < x_0 = 0 < x_1 < \dots < x_N$ .

<sup>25</sup> For any  $x \in \mathbb{R}$ ,  $[x]$  denotes its integer part.

Firstly, we want a discretization  $L$  of the operator  $\mathcal{A} - \rho \text{Id}$  such that  $-L$  is an SDD  $L_0$ -matrix as per (A1). A standard way to do this is to approximate the first (resp. second) order derivatives with forward and backward (resp. central) differences in such a way that we approximate the ordinary differential equation (ODE)  $\frac{1}{2}\sigma^2 V'' + \mu V' - \rho V + f = 0$  with an *upwind* (or *positive-coefficients*) scheme. More precisely, for each  $x = x_i \in G$  we approximate the first derivative with a forward (resp. backward) difference if its coefficient in the previous equation is nonnegative (resp. negative) in  $x_i$ ,

$$V'(x_i) \approx \frac{V(x_{i+1}) - V(x_i)}{x_{i+1} - x_i} \quad \text{if } \mu(x_i) \geq 0 \quad \text{and}$$

$$V'(x_i) \approx \frac{V(x_i) - V(x_{i+1}))}{x_i - x_{i+1}} \quad \text{if } \mu(x_i) < 0$$

and the second derivative by

$$V''(x_i) \approx \frac{V(x_{i+1}) - V(x_i)}{(x_{i+1} - x_i)(x_{i+1} - x_{i-1})} - \frac{V(x_i) - V(x_{i-1}))}{(x_i - x_{i-1})(x_{i+1} - x_{i-1})}.$$

In the case of an equispaced grid with step size  $h$ , this reduces to<sup>26</sup>

$$V'(x) \approx \frac{V(x + \text{sgn}(\mu(x))h) - V(x)}{\text{sgn}(\mu(x))h} \quad \text{and}$$

$$V''(x) \approx \frac{V(x + h) - 2V(x) + V(x - h)}{h^2}.$$

For the previous stencils to be defined in the extreme points of the grid, we consider two additional points  $x_{-N-1}, x_{N+1}$  and replace  $V(x_{-N-1}), V(x_{N+1})$  in the previous formulas by some values resulting from artificial boundary conditions. A common choice is to impose Neumann conditions to solve for  $V(x_{-N-1}), V(x_{N+1})$  using the first order differences from before. For example, in the equispaced grid case, given  $\text{LBC}, \text{RBC} \in \mathbb{R}$  we solve for  $V(x_{-N} - h)$  (resp.  $V(x_N + h)$ ) from the Neumann condition

$$\text{LBC} = V' \left( x_{-N} - h \mathbb{1}_{\{\mu(x_{-N}) \geq 0\}} \right) \quad (\text{resp. } \text{RBC} = V' \left( x_N + h \mathbb{1}_{\{\mu(x_N) < 0\}} \right)),$$

yielding  $V(x_{-N} - h) \approx V(x_{-N}) - \text{LBC}h$  (resp.  $V(x_N + h) \approx V(x_N) + \text{RBC}h$ ). The choice of  $\text{LBC}, \text{RBC}$  is problem-specific and intrinsically linked to that of  $x_N$ , although it does not affect the properties of the discrete operators. See more details in Sect. 3.

The described procedure leads to a discretization of the ODE as  $Lv + f = 0$ , with  $L$  satisfying the properties we wanted (with *strict* diagonal dominance due to  $\rho > 0$ ). Note that the values of  $f$  at  $x_{-N}, x_N$  need to be modified to account for the boundary conditions.

<sup>26</sup>  $\text{sgn}$  denotes the sign function,  $\text{sgn}(x) = 1$  if  $x \geq 0$  and  $-1$  otherwise.

**Remark 2.6.1** One could increase the overall order of approximation by using central differences as much as possible for the first order derivatives, provided the scheme remains upwind (see [25,38] for more details). This is not done here in order to simplify the presentation.

We now approximate the impulse constraint sets  $\mathcal{Z}(x)$  ( $x \in \mathbb{R}$ ) by finite sets  $\emptyset \neq Z(x) \subseteq [0, +\infty)$  ( $x \in G$ ), such that  $Z(x) = \{0\}$  if  $x \geq 0$ , and define the impulse operators

$$B(\delta)v(x) = v\llbracket x + \delta(x) \rrbracket, \text{ for } v \in \mathbb{R}^G, \delta \in Z, x \in G,$$

where  $v\llbracket y \rrbracket$  denotes linear interpolation of  $v$  on  $y$  using the closest nodes on the grid, and  $v\llbracket y \rrbracket = v(x_{\pm N})$  if  $\pm y > \pm x_{\pm N}$  (i.e., “no extrapolation”). This univocally defines the discrete loss and gain operators  $M$  and  $H$  as per (2.1.2), as well as the optimal impulse  $\delta^*$  according to (2.1.3). The set of discrete strategies  $\Phi$  is defined as in (2.2.1).

This general discretization scheme satisfies assumptions (A0)– (A2) and one can impose some regularity conditions on the sets  $\mathcal{Z}(x)$  and  $Z(x)$  such that the solutions of the discrete QVI problems (2.2.2) converge locally uniformly to the unique viscosity solution of the analytical impulse control problem, as the grid is refined.<sup>27</sup> See [2,6] for more details.

**Example 2.6.2** In the case where  $\mathcal{Z}(x) = [0, +\infty)$  for  $x < 0$ , a natural and most simple choice for  $Z(x)$  is  $Z(x_i) = \{0, x_{i+1} - x_i, \dots, x_N - x_i\}$  for  $i < 0$ . In this case,  $B(\delta)v(x) = v(x + \delta(x))$  and  $Hv(x) = v(x - \delta^*(-x)) + g(x, -\delta^*(-x))$ . This choice, however, does not satisfy (A1’).

In order to preclude infinite simultaneous interventions it is enough to constrain the size of the impulses so that the symmetric point of the grid cannot be reached. That is,  $Z(x) \subseteq [0, -2x)$  for any  $x \in G_{<0}$ . In this case, the scheme satisfies the stronger conditions (A0’), (A1), (A2) (and in particular, (A1’)). Intuitively, each positive impulse will lead to a state which is at least one node closer to  $x_0 = 0$ , where no player intervenes. Practically, it makes sense to make this choice when one suspects (or wants to check whether) there is a symmetric NE with no “far-reaching impulses”, in the previous sense.

**Proof** To check that (A0’’) is indeed satisfied, consider some arbitrary  $\varphi = (I, \delta)$ ,  $\bar{\varphi} = (\bar{I}, \bar{\delta})$  and  $x_i \in G$ . Note first that if  $x_i$  belongs to the common continuation region  $\bar{I}^c \cap (-I)^c$  (as it is the case for  $x_0 = 0$ ), then the  $i$ -th row of  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is equal to the  $i$ -th row of  $-L$ , which is SDD by (A1). We claim that, in general, either the  $i$ -th row of  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is SDD or there is an edge in its graph from  $x_i$  to some  $x_j$  with  $|j| < |i|$ . This immediately implies that (in the notation of (A0’)) one can take  $m = N$ , as there will always be a walk in graph  $((\mathbb{A} - \mathbb{B})(\bar{\varphi}^k, \bar{\varphi}^{k+1}))$  from  $x_i$  to some SDD row, with length at most  $N$ . (The longest that might need to be walked is all the way until reaching the 0-th row, with the distance decreasing in one unit per step.)

<sup>27</sup> Additional technical conditions include costs bounded away from zero and a comparison principle for the analytical QVI.

To prove the claim, suppose first that  $x_i \in \bar{I}$  (i.e., the intervention region of the player). Hence,  $i < 0$ . If  $\bar{\delta} = 0$ , then the  $i$ -th row of  $(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})$  is again equal to the  $i$ -th row of  $\text{Id}$ , thus SDD. If  $0 < \bar{\delta} < -2x_i$ , then there exist an index  $k$  and  $\alpha, \beta \geq 0$  such that  $i < k \leq |i|$ ,  $\alpha + \beta = 1$  and  $x_i + \bar{\delta} = \alpha x_{k-1} + \beta x_k$ . Accordingly,

$$[(\mathbb{A} - \mathbb{B})(\varphi, \bar{\varphi})]_{ij} = \mathbb{1}_{\{i=j\}} - \alpha \mathbb{1}_{\{j=k-1\}} - \beta \mathbb{1}_{\{j=k\}}, \quad \text{for all } j.$$

If  $x_i < x_i + \bar{\delta} \leq x_{|i-1}$ , then we can assume without loss of generality that  $\beta > 0$  and take  $j = k$ . (If  $\beta = 0$ , simply relabel  $k - 1$  as  $k$ .) If  $x_{|i-1} < x_i + \bar{\delta} < -x_i = x_{|i|}$ , then  $\alpha > 0$  and we can take  $j = k - 1$ .

The case of  $x_i \in -I$  is symmetric and can be proved in the same manner. □

**Example 2.6.3** If  $\mathcal{Z}(x) = [0, +\infty)$  for  $x < 0$ , the analogous of Example 2.6.2 is now  $Z(x_i) = \{0, x_{i+1} - x_i, \dots, x_{-i-1} - x_i\}$  for  $i < 0$ .

**Remark 2.6.4** Consider Example 2.6.3 in the context of Remark 2.1.1. As in Proposition 2.5.6 and due to Theorem A.1.7, the less impulses needed between the two players to reach the common continuation region, the faster that the composition of the fixed-point operators of Algorithm 2.2.1 becomes contractive. Hence, one could intuitively expect that when close enough to the solution, the choice of the maximum arg-maximum in (2.1.3) improves the performance of Algorithm 2.2.1. This is another motivation for such choice.

### 3 Numerical Results

This section presents numerical results obtained on a series of experiments. See Introduction and Sect. 1.3 for the motivation and applications behind some of them. We do not assume additional constraints on the impulses in the analytical problem. All the results presented were obtained on equispaced grids with step size  $h > 0$  (to be specified) and with a discretization scheme as in Sect. 2.6 and Example 2.6.3. The extreme points of the grid are displayed on each graph.

For the games with linear costs and gains of the form  $c(x, \delta) = c^0 + c^1 \delta$  and  $g(x, \delta) = g^0 + g^1 \delta$ , with  $c^0, c^1, g^0, g^1$  constant, the artificial boundary conditions were taken as  $\text{LBC} = c^1$  and  $\text{RBC} = g^1$  for a sufficiently extensive grid. They result from the observation that on a hypothetical symmetric NE of the form  $\varphi^* = ((-\infty, \bar{x}], \delta^*(x) = y^* - x)$ , with  $\bar{x} < 0$ ,  $\bar{x} < y^* \in \mathbb{R}$ , the equilibrium payoff verifies  $V(x) = V(y^*) - c^0 - c^1(y^* - x)$  for  $x < \bar{x}$  and  $V(x) = V(-y^*) + g^0 + g^1(x + y^*)$  for  $x > -\bar{x}$ . For other examples, LBC, RBC and the grid extension were chosen by heuristic guesses and/or trial and error. However, in all the examples presented the error propagation from poorly chosen LBC,RBC was minimal.

The initial guess was set as  $v_0 = 0$  and its induced strategy in all cases. SOLVEIMPULSECONTROL was chosen as Sect. 2.3.1 (see Remark 2.3.4) with  $\widetilde{\text{tol}} = 10^{-15}$  and  $\lambda = 1$ .<sup>28</sup> Its convergence was exact however, in all the examples. Instead of fixing

<sup>28</sup> For very fine grids, one should increase the value of  $\widetilde{\text{tol}}$  to avoid stagnation (see, e.g, Remark 2.2.4).

a terminal tolerance  $tol$  beforehand, we display the highest accuracy attained and number of iterations needed.

Section 3.1 considers a fixed grid and games where the results point to the existence of a symmetric NE as per Corollary 1.3.9. Not having an analytical solution to compare with, results are assessed by means of the percentage difference between the iterates

$$\text{Diff} := \left\| (v^{k+1} - v^k) / \max\{|v^{k+1}|, scale\} \right\|_{\infty},$$

with  $scale = 1$  as in [4], and the maximum pointwise residual to the system of QVIs (2.1.1), defined for  $v \in \mathbb{R}^G$  by setting  $I = \{Lv + f \leq Mv - v\} \cap G_{<0}$ ,  $C = I^c$  and

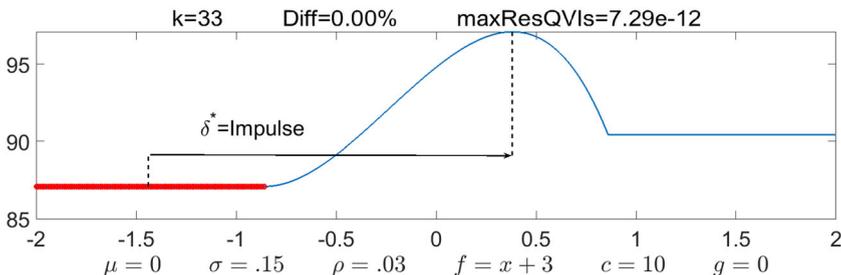
$$\text{maxResQVIs}(v) := \|\max\{Lv + f, Mv - v\} \mathbb{1}_{-C} + (Hv - v) \mathbb{1}_{-I}\|_{\infty}.$$

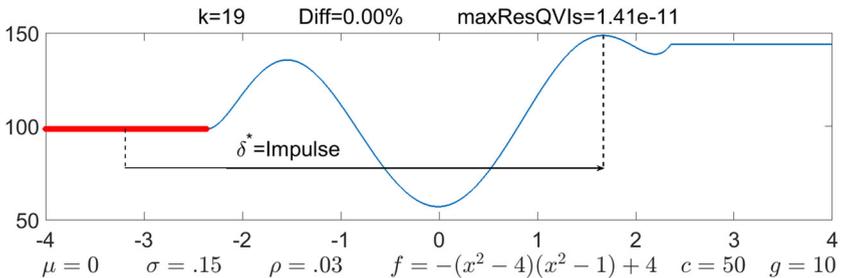
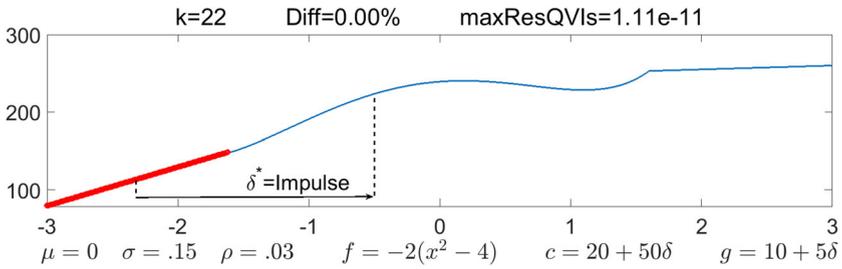
Section 3.2 considers the only symmetric games in the literature with (semi-) analytical solution: the central bank linear game [1] and the cash management game [10], and computes the errors made by discrete approximations, showing in particular the effect of refining the grid. Not considered here is the strategic pollution control game [24], due to its inherent non-symmetric nature. Section 3.3 comments on results obtained for games without NEs. Finally, Sect. 3.4 shows results that go beyond the scope of the currently available theory for impulse games.

### 3.1 Convergence to Discrete Solution on a Fixed Grid

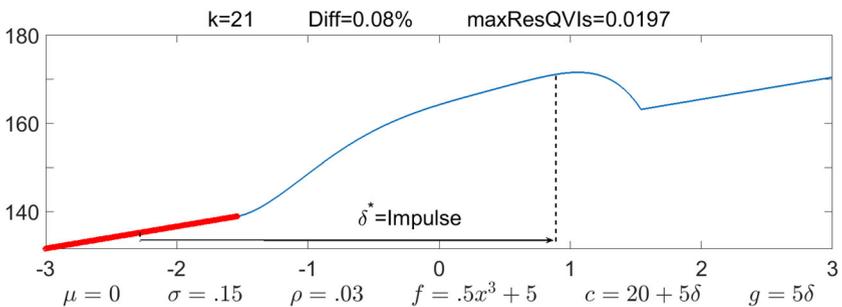
Throughout this section the grid step size is fixed as  $h = 0.01$ , unless otherwise stated (although results were corroborated by further refinements). Each figure specifies the structure,  $\mathcal{G} = (\mu, \sigma, \rho, f, c, g)$ , of the symmetric game solved and shows the numerical solutions at the terminal iteration for the equilibrium payoff,  $v^k$ , and NE. Graphs plot payoff versus state of the process. The intervention region is displayed in red over the graph of the payoff for presentation purposes.

We focus on games with higher costs than gains, as the opposite typically leads to players attempting to apply infinite simultaneous impulses [1] (i.e., inducing a gain from the opponent’s intervention is “cheap”) leading to degenerate games. The following games resulted in exact convergence in finite iterations, which guarantees a solution of (2.1.1) was reached (Corollary 2.5.3), although very small acceptable errors were reached sooner.

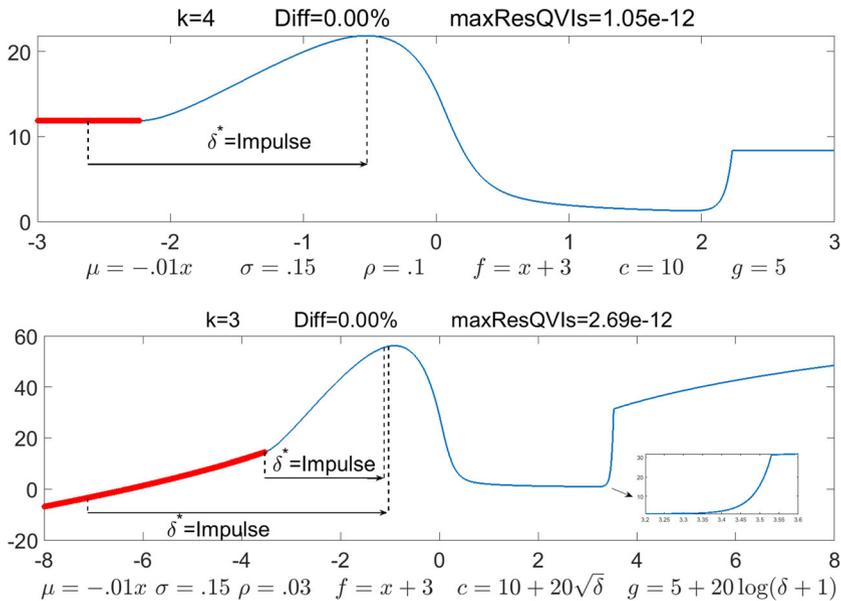




The following is an example in which the accuracy stagnates. At that point, the iterates start going back and forth between a few values. Although we cannot guarantee that we are close to a solution of (2.1.1), the results seem quite convincing, with both Diff and maxResQVIs reasonably low. In fact, simply halving the step size to  $h = 0.005$  produces a substantial improvement of Diff=9.16e-11% and maxResQVIs=9.19e-11 in  $k = 33$  iterations.



The previous games have a state variable evolving as a scaled Brownian motion. We now move on to a mean reverting OU process with zero long term mean. (Recall that any other value can be handled simply by shifting the game.) In general, the experiments with these dynamics converged exactly in a very small number of iterations.



Note how all the games treated in this section exhibit a typical feature known to hold for simpler symmetric games [1,3,10]: the equilibrium payoff of the player is only  $C^1$  at the border of the intervention region  $\partial\mathcal{I}^* = \{\mathcal{AV} - \rho V + f = \mathcal{MV} - V\} \cap \mathbb{R}_{<0}$ , and only continuous at the border of the opponent’s intervention region  $-\partial\mathcal{I}^*$ . In floating point arithmetic, the former makes the discrete approximation of  $\partial\mathcal{I}^*$  particularly elusive, while the latter can lead to high errors when close to  $-\partial\mathcal{I}^*$ . As a consequence, Sect. 2.3.1 (or any equivalent) will often misplace a few grid nodes between intervention and continuation regions, which will in turn make the residual resQVIs “spike” on the opponent’s side. Thus, a large value of maxResQVIs can at times be misleading, and further inspection of the pointwise residuals is advisable.

As a matter of fact, halving the step size to  $h = 0.005$  in the last example results in exact convergence with terminal maxResQVIs= 2510, but the residuals on all grid nodes other than the “border” of the opponent’s intervention region and a contiguous one are less than  $1.43e-11$ . This is an extreme example propitiated by the almost vertical shape of the solution close to such border. Thus, while it is useful in practice to consider a stopping criteria for Algorithm 2.2.1 based on maxResQVIs, this phenomenon needs to be minded.

The last example also shows how impulses at a NE can lead to different endpoints depending on the state of the process. This is often the case when costs are nonlinear. In fact:

**Lemma 3.1.1** *Let  $(\mathcal{I}^*, \delta^*)$  and  $V$  be a symmetric NE and equilibrium payoff as in Corollary 1.3.9. Assume that  $\mathcal{Z}(x) = [0, +\infty)$  for  $x < 0$  and  $c = c(x, \delta) \in C^2(\mathbb{R} \times (0, +\infty))$ , and consider the re-parametrization  $\bar{c} = \bar{c}(x, y) := c(x, y - x)$ . Suppose that  $y^* := x + \delta^*(x)$  is constant for all  $x \in (\mathcal{I}^*)^\circ$ .<sup>29</sup> Then  $\bar{c}_{xy}(x, y^*) \equiv 0$  on  $(\mathcal{I}^*)^\circ$ .*

<sup>29</sup>  $A^\circ$  denotes the interior of the set  $A \subseteq \mathbb{R}$ .

The result is immediate since  $\mathcal{MV}(x) = \sup_{y>x} \{V(y) - \bar{c}(x, y)\} = V(y^*) - \bar{c}(x, y^*)$  on  $(\mathcal{I}^*)^\circ$  and  $V \in C^1(-\mathcal{C}^*)$ . ( $y^* \notin -\mathcal{I}^*$  or there would be infinite simultaneous impulses.) While the sufficient condition  $\bar{c}_{xy}(x, y^*) \equiv 0$  for some  $y^*$  is verified for linear costs, it is not in general and certainly not for  $c(x, \delta) = 10 + 20\sqrt{\delta}$  as in the last example.

### 3.2 Convergence to Analytical Solution with Refining Grids

A convergence analysis from discrete to analytical solution with refining grids is outside the scope of this paper, and seems to be far too challenging when a viscosity solutions framework is yet to be developed. Instead, we present here a numerical validation using the solutions of the linear and cash management games. We focus first and foremost in the former, as it has an almost fully analytical solution, with only one parameter to be found numerically as opposed to four for the latter. The structure of the game is defined with parameter values used in [1] (also in [3]). To minimize rounding errors from floating point arithmetic, we proceed as in [4] considering grids made up entirely of machine numbers.<sup>30</sup> The results are displayed in Table 1.

For each step size, Algorithm 2.2.1 either converged or was terminated upon stagnation. Regardless, we can see the errors made when approximating the analytical solution are quite satisfactory in all cases, and overall decrease as  $h \rightarrow 0$ . Moreover, we see once again how the “spiking” of the residual can be misleading (the highest value was always at the “border” of the opponent’s intervention region).

An exact symmetric NE for this game (up to five significant figures) is given by the intervention region  $\mathcal{I}^* = (-\infty, -2.8238]$  and impulse function  $\delta^*(x) = 1.5243 - x$ , while the approximation given by Algorithm 2.2.1 with  $h = 1/64$  is  $((-\infty, -2.8125], 1.5313 - x)$ , with absolute errors on the parameters of no more than the step size.

The cash management game [10] with unidirectional impulses can be embedded in our framework by reducing its dimension with the change of variables  $x = x_1 - x_2$ , changing minimization by maximization and relabelling the players. With the parameter values of [10, Fig.1b], it translates into the game:  $\mu = 0, \sigma = 1, \rho = .5, f = -|x|, c = 3 + \delta, g = -1$ . The authors found numerically a symmetric NE approximately equal to  $((-\infty, -5.658], 0.686 - x)$ , while Algorithm 2.2.1 with  $x_N = 8$  and  $h = 1/64$  gives  $((-\infty, -5.6563], 0.6875 - x)$ . The absolute difference on the parameters is once again below the grid step size.

### 3.3 Games Without Nash Equilibria

It is natural to ask how Algorithm 2.2.1 behaves on games without symmetric NEs, and whether anything can be inferred from the results. For the linear game, two cases without NEs (symmetric or not) are addressed in [1]: “no fixed cost” and “gain greater than cost”. Both of them yield degenerate “equilibria” where the players perform infi-

<sup>30</sup> By *machine numbers* we mean those that can be represented exactly in IEEE standard floating-point number system.

**Table 1** Convergence to analytical solution when refining equispaced symmetric grid with step size  $h$  and endpoint  $x_N = 4$ 

$h$	%error	Its.	maxResQVIs
1	6.67%	17	8.88e-16
1/2	8.33%	13	5.33e-15
1/4	0.23%	4	13.2
1/8	0.21%	8	15.1
1/16	0.16%	8	30
1/32	0.07%	21	21.2
1/64	0.0043%	37	0.343

Game:  $\mu = 0$ ,  $\sigma = .15$ ,  $\rho = .02$ ,  $f = x + 3$ ,  $c = 100 + 15\delta$ ,  $g = 15\delta$ . %error :=  $\|(v - V)/V\|_\infty$ , with  $V$  exact solution and  $v$  discrete approximation after Its iterations.  $\|\cdot\|_\infty$  is computed over the grid

nite simultaneous interventions. When tested for several parameters, Algorithm 2.2.1 converged in finitely many iterations (although rather slowly) and yielded the exact same type of “equilibria”.<sup>31</sup> For a fine enough grid, the previous can be identified heuristically by some node in the intervention region that would be shifted to its symmetric one (*infinite alternated interventions*), or to its immediate successor over and over again, until reaching the continuation region (*infinite one-sided interventions*).<sup>32</sup>

In the first case, we recovered the limit “equilibrium payoffs” of [1, Props.4.10,4.11]. Intuitively, the players in this game take advantage of free interventions, whether by no cost or perfect compensation, in order to shift the process as desired. Note that when  $c = g$ , the impulses that maximize the net payoff are not unique.

In the second case, grid refinements showed the discrete payoffs to diverge towards infinity at every point. This is again consistent with the theory: each player forces the other one to act, producing a positive instantaneous profit. Iterating this procedure infinitely often leads to infinite payoffs for every state.

Tested games in which Algorithm 2.2.1 failed to converge (and not due to stagnation nor a poor choice of the grid extension) were characterized by iterates reaching a cycle, typically with high values of Diff and maxResQVIs regardless of the grid step size. In many cases, the cycles would visit at least one payoff inducing infinite simultaneous impulses. While this might, potentially and heuristically, be indicative of the game admitting no symmetric NE, there is not much more than can be said at this stage.

### 3.4 Beyond the Verification Theorem

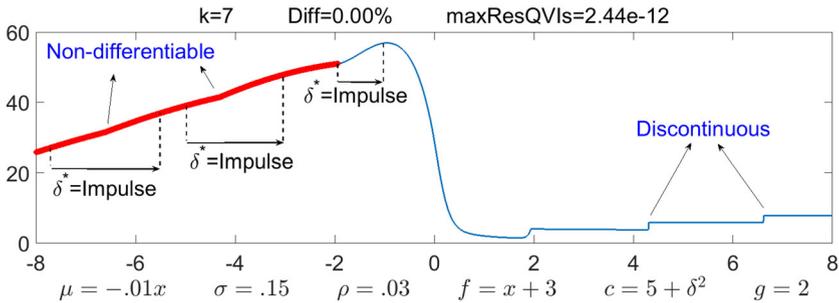
We now present two games in which the solution of the discrete QVIs system (2.1.1) found with Algorithm 2.2.1 (by exact convergence) does not comply with the continuity and smoothness assumptions of Corollary 1.3.9. However, the results are sensible enough to heuristically argue they may correspond to NEs beyond the scope of the

<sup>31</sup> There were also cases of stagnation, improved by refining the grid as in Sect. 3.1.

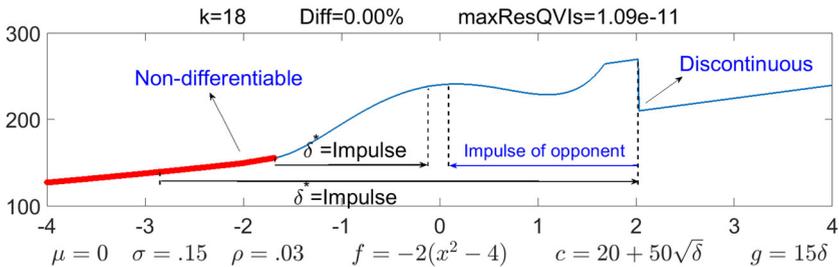
<sup>32</sup> More precisely, due to our choice of  $Z(x)$  a node can be shifted at most to that immediately preceding its symmetric one.

Verification Theorem. In both cases  $h = 0.01$ . Finer grids yielded the same qualitative results.

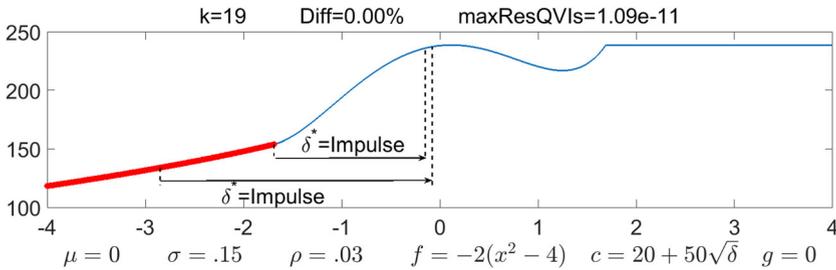
The first game considers costs convex in the impulse. When far enough from her continuation region, it is cheaper for the player to apply several (finitely many) simultaneous impulses instead of one, to reach the state she wishes to (cf. Remark 1.2.4). In this game, the optimal impulse  $\delta^*$  becomes discontinuous, and its discontinuity points are those in the intervention region where the equilibrium payoff is non-differentiable. These, in turn, translate into discontinuities on the opponent's intervention region (one degree of smoothness less, as with the border of the regions).



The second game considers linear gains, quadratic running payoffs and costs concave on the impulse. The latter makes the player shift the process towards her continuation region (cf. Remark 1.2.4). However, when far enough from the border, instead of shifting the process directly to her “preferred area”, the player chooses to pay a bit more to force her opponent’s intervention, inducing a gain and letting the latter pay for the final move. Once again, this causes  $\delta^*$  to be discontinuous and leads to a non-differentiable (resp. discontinuous) point for the equilibrium payoff in the intervention (resp. opponent’s intervention) region.



Under the previous reasoning, one could intuitively guess that setting  $g = 0$  in this game should remove the main incentive the player has to force her opponent to act. This is in fact the case, as shown below. As a result,  $\delta^*$  becomes continuous and the equilibrium payoff falls back into the domain of the Verification Theorem.



We remark that, should the previous solutions correspond indeed to NEs, then the alternative semi-analytical approach of [24] could not have produced them either, as that method can only yield continuous equilibrium payoffs.

#### 4 Concluding Remarks

This paper presents a fixed-point policy-iteration-type algorithm to solve systems of quasi-variational inequalities resulting from a Verification Theorem for symmetric nonzero-sum stochastic impulse games. In this context, the method substantially improves the only algorithm available in the literature while providing the convergence analysis that we were missing. Graph-theoretic assumptions relating to weakly chained diagonally dominant matrices, which naturally parallel the admissibility of the players strategies, allow to prove properties of contractiveness, boundedness of iterates and convergence to solutions. A result of theoretical interest giving sufficient conditions for convergence is also proved. Additionally, a novel provably convergent impulse control solver is provided.

Equilibrium payoffs and Nash equilibria of games too challenging for the available analytical approaches are computed with high precision on a discrete setting. Numerical validation with analytical solutions is performed when possible, with reassuring results, but it is noted that grid refinements may be needed at times to overcome stagnation. Thus, formalising the approximating properties of the discrete solutions as well as deriving stronger convergence results for the algorithm may need a viscosity solutions framework currently missing in the theory. This is further substantiated by the irregularity of the solutions, particularly those found which escape the available theoretical results. This motivates further research while providing a tool that can effectively be used to gain insight into these very challenging problems.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## A Matrix and Graph-Theoretic Definitions and Results

For the reader’s convenience, this appendix summarizes some important algebraic and graph-theoretic definitions and results used throughout the paper. More details can be found in the references given below. Henceforth,  $A \in \mathbb{R}^{N \times N}$  is a real matrix,  $\text{Id} \in \mathbb{R}^{N \times N}$  is the identity,  $\rho(\cdot)$  denotes the spectral radius and  $\mathbb{R}^N, \mathbb{R}^{N \times N}$  are equipped with the elementwise order. We talk about rows and “states” interchangeably.

- Definitions A.1.** (D1)  $A$  is a *Z-matrix* if it has nonpositive off-diagonal elements.  
 (D2)  $A$  is an *L-matrix* (resp. *L<sub>0</sub>-matrix*) if it is a Z-matrix with positive (resp. non-negative) diagonal elements.  
 (D3)  $A$  is an *M-matrix* if  $A = s \text{Id} - B$  for some matrix  $B \geq 0$  and scalar  $s \geq \rho(B)$ .  
 (D4)  $A$  is *monotone* if it is nonsingular and  $A^{-1} \geq 0$ . Equivalently,  $A$  is monotone if  $Ax \geq 0$  implies  $x \geq 0$  for any  $x \in \mathbb{R}^N$ .  
 (D5) The  $i$ -th row of  $A$  is *weakly diagonally dominant (WDD)* (resp. *strictly diagonally dominant or SDD*) if  $|A_{ii}| \geq \sum_{j \neq i} |A_{ij}|$  (resp.  $>$ ).  
 (D6)  $A$  is *WDD* (resp. *SDD*) if every row of  $A$  is WDD (resp. SDD).  
 (D7) The *directed graph* of  $A$  is the pair  $\text{graph}A := (V, E)$ , where  $V := \{1, \dots, N\}$  is the set of *vertexes* and  $E \subseteq V \times V$  is the set of *edges*, such that  $(i, j) \in E$  iff  $A_{ij} \neq 0$ .  
 (D8) A *walk*  $\omega$  in  $\text{graph}A = (V, E)$  from vertex  $i$  to vertex  $j$  is a nonempty finite sequence  $(i, i_1), (i_1, i_2), \dots, (i_{k-1}, j) \subseteq E$ , which we denote by  $\omega : i \rightarrow i_1 \rightarrow \dots \rightarrow i_{k-1} \rightarrow j$ .  $|\omega| := k$  is called the *length* of the walk  $\omega$ .  
 (D9)  $A$  is *weakly chained diagonally dominant (WCDD)* if it is WDD and for each WDD row of  $A$  there is a walk in  $\text{graph}A$  to an SDD row (identifying vertexes and rows).  
 (D10)  $A$  is (right) *substochastic or sub-Markov* (resp. *stochastic or Markov*) if  $A \geq 0$  and each row sums at most one (resp. exactly one). Equivalently,  $A$  is substochastic if  $A \geq 0$  and  $\|A\|_\infty \leq 1$ . (Recall that  $\|A\|_\infty$  is the maximum row-sum of absolute values.)  
 (D11) If  $A$  is a WDD (resp. substochastic) matrix, its set of “non-trouble states” (or rows) is  $J[A] := \{i : \text{the } i\text{-th row of } A \text{ is SDD}\}$  (resp.  $\hat{J}[A] := \{i : \sum_j A_{ij} < 1\}$ ). For each  $i$ , we write  $P_i[A] := \{\text{walks in } \text{graph}A \text{ from } i \text{ to some } j \in J[A]\}$  (resp. we define  $\hat{P}_i[A]$  analogously). The *index of connectivity* (resp. *index of contraction*) of  $A$  [7] is<sup>33</sup>

$$\text{con}A := \left( \sup_{i \notin J[A]} \left\{ \inf_{\omega \in P_i[A]} |\omega| \right\} \right)^+ \quad (\text{resp. we define } \widehat{\text{con}}A \text{ analogously}).$$

It is clear that  $\text{SDD} \implies \text{WCDD} \implies \text{WDD}$ , and by definition,  $\text{L-matrix} \implies \text{L}_0\text{-matrix} \implies \text{Z-matrix}$ . Also by definition, if  $A$  is WDD then:  $A$  is WCDD  $\iff \text{con}A < +\infty$ .

**Proposition A.1.2** (e.g., [36] or [4, Lem.3.2]) *Any WCDD matrix is nonsingular.*

<sup>33</sup>  $(\cdot)^+$  denotes positive part,  $\inf \emptyset = +\infty$  and  $\sup \emptyset = -\infty$ . The index is the least length that needs to be walked on  $\text{graph}A$  to reach the non-trouble states when starting from an arbitrary trouble one.

**Proposition A.1.3** (e.g., [7, Prop.2.15 and 2.17])

*Nonsingular M-matrix*  $\iff$  *monotone L-matrix*  $\iff$  *monotone Z-matrix*.

**Theorem A.1.4** (e.g., [7, Thm.2.24]) *WCDD L<sub>0</sub>-matrix*  $\iff$  *WDD nonsingular M-matrix*.<sup>34</sup>

**Proposition A.1.5** (see proof of [7, Lem.2.22]) *A is substochastic if and only if Id – A is a WDD L<sub>0</sub>-matrix and A has non-negative diagonal elements. In such case,  $\hat{J}[A] = J[\text{Id} – A]$ , they have the same directed graphs (except possibly for self-loops  $i \rightarrow i$ ) and  $\widehat{\text{con}}A = \text{con}[\text{Id} – A]$ .*

For the following theorem, recall the characterization of the spectral radius  $\rho(A) = \inf\{\|A\| : \|\cdot\| \text{ is a matrix norm}\}$  and Gelfand’s formula  $\rho(A) = \lim_{n \rightarrow +\infty} \|A^n\|^{1/n}$ , for any matrix norm  $\|\cdot\|$ . Note also that if A is substochastic, then  $A^n$  is also substochastic for any  $n \in \mathbb{N}_0$ ,  $\|A^n\|_\infty \leq 1$  and  $\rho(A) \leq 1$ .

**Theorem A.1.6** ([7, Thm.2.5 and Cor.2.6]) *Suppose A is substochastic. Then*

$$\widehat{\text{con}}A = \inf\{n \in \mathbb{N}_0 : \|A^{n+1}\|_\infty < 1\}.$$

*In particular,  $\widehat{\text{con}}A < +\infty$  if and only if  $\rho(A) < 1$ .*

The indices of contraction and connectivity can be generalized in a natural way to sequences  $(A_k) \subseteq \mathbb{R}^{N \times N}$  by considering walks  $i_1 \rightarrow i_2 \rightarrow \dots$  such that  $i_k \rightarrow i_{k+1}$  is an edge in  $\text{graph}A_k$  (see [7, App.B] for more details). Theorem A.1.6 extends in the following way:

**Theorem A.1.7** ([7, Thm.B.2]) *Suppose  $(A_k)$  are substochastic matrices and consider the sequence of products  $(B_k)$ , where  $B_k := A_1 \dots A_k$ . Then,*

$$\widehat{\text{con}}[(A_k)] = \inf\{k \in \mathbb{N}_0 : \|B_{k+1}\|_\infty < 1\}.$$

*Equivalently,  $1 = \|B_0\|_\infty = \dots = \|B_\alpha\|_\infty > \|B_{\alpha+1}\|_\infty \geq \dots$ , for  $\alpha = \widehat{\text{con}}[(A_k)]$ .*<sup>35</sup>

**Proof of Lemma 2.4.4** For brevity, we omit the dependence on  $\varphi, \bar{\varphi}$  from the notation.

It follows from Proposition 2.4.2 that  $\mathbb{A}^{-1}\mathbb{B} \geq 0$ , since  $\mathbb{B} \geq 0$  and  $\mathbb{A}$  is monotone. To see that its rows sum up to one, let  $\mathbf{1} \in \mathbb{R}^G$  be the vector of ones. It is easy to check using  $\mathbb{A}$  and  $\mathbb{B}$  explicit expressions and (A1)–(A2), that  $\mathbb{A}\mathbb{A}^{-1}\mathbb{B}\mathbf{1} = \mathbb{B}\mathbf{1} \leq \mathbb{A}\mathbf{1}$ , which implies  $\mathbb{A}^{-1}\mathbb{B}\mathbf{1} \leq \mathbf{1}$ . This proves that  $\mathbb{A}^{-1}\mathbb{B}$  is substochastic.

The WDD L<sub>0</sub>-property of  $\mathbb{A} – \mathbb{B}$  is due to (A1). To verify this, note first that by straightforward computations, if a matrix  $A_0$  has the previous property then  $SA_0S$  must also have it. Then simply observe that the  $i$ -th row of  $\mathbb{A} – \mathbb{B}$  is equal to the  $i$ -th row of  $\text{Id} – B(\bar{\delta})$  (if  $x_i \in \bar{I}$ ), or  $–L$  (if  $x_i \in \bar{I}^c \cap (–I)^c$ ) or  $\text{Id} – SB(\delta)S = S(\text{Id} – B(\delta))S$  (if  $x_i \in –I$ ), all of which are WDD L<sub>0</sub>-matrices by (A1).

<sup>34</sup> [7, Thm.2.24] is formulated in terms of L-matrices instead. However, it is trivial to see that: WCDD L<sub>0</sub>-matrix  $\iff$  WCDD L-matrix.

<sup>35</sup> The equivalence is due to  $(A_k)$  being substochastic matrices, which implies that  $\|B_k\|_\infty \leq 1$  for all  $k$ . Note that the case  $\alpha = +\infty$  reads  $\|B_k\|_\infty = 1$  for all  $k$ .

We prove now the last statement,  $\widehat{\text{con}}[\mathbb{A}^{-1}\mathbb{B}] \leq \text{con}[\mathbb{A} - \mathbb{B}]$ . Note that by Proposition A.1.5,  $\mathbb{A}^{-1}\mathbb{B}$  and  $\text{Id} - \mathbb{A}^{-1}\mathbb{B}$  share the same sets of non-trouble rows, directed graphs (except possibly for self loops) and  $\widehat{\text{con}}[\mathbb{A}^{-1}\mathbb{B}] = \text{con}[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]$ . Thus, we want see that for any non-SDD row of  $\text{Id} - \mathbb{A}^{-1}\mathbb{B}$ , if an SDD row of  $\mathbb{A} - \mathbb{B}$  can be reached by a walk in the graph of the latter, then there must be a shorter walk to an SDD row in the graph of the former. We proceed by showing several results:

(i) *There is a walk from  $x_i$  to  $x_j$  in  $\text{graph}\mathbb{A}$  if and only if  $\mathbb{A}_{ij}^{-1} > 0$ .<sup>36</sup> In particular,  $\mathbb{A}^{-1}$  has strictly positive diagonal elements. (More in general, both statements are true for any WCDD  $L_0$ -matrix, with the same proof)*

Recall first that  $\mathbb{A}$  is a WCDD  $L_0$ -matrix and so it must have positive diagonal elements, or there would be a null row disconnected from every SDD row. Thus,  $x_i \rightarrow x_i$  is always in  $\text{graph}\mathbb{A}$  and we must show that  $\mathbb{A}_{ii}^{-1} > 0$ . If  $\mathbb{A}_{ii}^{-1} = 0$ , then by monotonicity and  $L_0$ -property, it would be  $1 = [\mathbb{A}\mathbb{A}^{-1}]_{ii} = \sum_{j \neq i} \mathbb{A}_{ij}\mathbb{A}_{ji}^{-1} \leq 0$ .

Suppose now  $i \neq j$  and that  $x_i$  is connected to  $x_j$  by a walk in  $\text{graph}\mathbb{A}$ . We proceed by induction in the length of the walk. If the length is one, suppose by contradiction that  $\mathbb{A}_{ij}^{-1} = 0$ . Then  $0 = [\mathbb{A}\mathbb{A}^{-1}]_{ij} = \sum_{k \neq i} \mathbb{A}_{ik}\mathbb{A}_{kj}^{-1} \leq \mathbb{A}_{ij}\mathbb{A}_{jj}^{-1} < 0$  by assumption and the diagonal case. For an arbitrary walk  $x_i \rightarrow y_1 \rightarrow \dots \rightarrow y_n = x_j$ , assume without loss of generality that it has no closed subwalks. Then  $y_1 = x_h$  for some  $h \neq i, j$ . Once again if  $\mathbb{A}_{ij}^{-1} = 0$ , this would give  $0 = [\mathbb{A}\mathbb{A}^{-1}]_{ij} = \sum_{k \neq i} \mathbb{A}_{ik}\mathbb{A}_{kj}^{-1} \leq \mathbb{A}_{ih}\mathbb{A}_{hj}^{-1} < 0$  by induction.

For the reciprocal, assume now that  $j \neq i$  and  $\mathbb{A}_{ij}^{-1} > 0$ . By the adjoint formula of the inverse and Leibniz (or permutations) formula of the determinant, there must exist some permutation  $\sigma : \{-N, \dots, N\} \setminus \{j\} \rightarrow \{-N, \dots, N\} \setminus \{i\}$  such that  $\prod_{k \neq j} \mathbb{A}_{k\sigma(k)}^{-1} \neq 0$ . Hence,  $x_i \rightarrow x_{\sigma(i)} \rightarrow x_{\sigma^2(i)} \rightarrow \dots \rightarrow x_j$  is a walk in  $\text{graph}\mathbb{A}$ .

(ii)  *$\text{Id} - \mathbb{A}^{-1}\mathbb{B}$  has no less SDD rows than  $\mathbb{A} - \mathbb{B}$ . Further,  $J[\mathbb{A} - \mathbb{B}] \subseteq J[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]$ :*

Consider some  $x_i \in J[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]^c$ , i.e.,  $\sum_j [\text{Id} - \mathbb{A}^{-1}\mathbb{B}]_{ij} = 0$ . We want to see that  $\sum_j [\mathbb{A} - \mathbb{B}]_{ij} = 0$ . We have  $0 = \sum_j [\text{Id} - \mathbb{A}^{-1}\mathbb{B}]_{ij} = \sum_j [\mathbb{A}^{-1}(\mathbb{A} - \mathbb{B})]_{ij} = \sum_k \mathbb{A}_{ik}^{-1} \sum_j [\mathbb{A} - \mathbb{B}]_{kj}$ . Since  $\mathbb{A}_{ik}^{-1}$  and  $\sum_j [\mathbb{A} - \mathbb{B}]_{kj}$  are non-negative for all  $k$  (monotonicity and  $L_0$  WDD property, respectively), one of the two must be zero for each  $k$ . But  $\mathbb{A}_{ii}^{-1} > 0$  by (i), giving what we wanted.

(iii) *If  $x_i \in -I$  and there is an edge  $x_i \rightarrow x_j$  in  $\text{graph}(\mathbb{A} - \mathbb{B})$  for some  $j \neq i$ , then this is also an edge in  $\text{graph}(\text{Id} - \mathbb{A}^{-1}\mathbb{B})$ :*

$[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]_{ij} = -[\mathbb{A}^{-1}\mathbb{B}]_{ij} = -\sum_k \mathbb{A}_{ik}^{-1}\mathbb{B}_{kj} \leq -\mathbb{A}_{ii}^{-1}\mathbb{B}_{ij} = \mathbb{A}_{ii}^{-1}[\mathbb{A} - \mathbb{B}]_{ij} < 0$ , where the last equality is due to  $x_i \in -I$  and  $j \neq i$ , and the strict inequality is by assumption and (i).

(iv) *If there is a walk in  $\text{graph}(\mathbb{A} - \mathbb{B})$ ,  $x_i = y_0 \rightarrow \dots \rightarrow y_n \rightarrow x_j \in J[\mathbb{A} - \mathbb{B}]$ , with  $y_m \in \bar{I}$  for all  $m$ , then  $x_i \in J[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]$ :*

Since  $\mathbb{A} - \mathbb{B}$  and  $\mathbb{A}$  agree for every row indexed in  $\bar{I}$ ,  $\mathbb{A}_{ij}^{-1} > 0$  by (i). It follows that  $\sum_k [\text{Id} - \mathbb{A}^{-1}\mathbb{B}]_{ik} = \sum_k [\mathbb{A}^{-1}(\mathbb{A} - \mathbb{B})]_{ik} = \sum_h \mathbb{A}_{ih}^{-1} \sum_k [\mathbb{A} - \mathbb{B}]_{hk} \geq \mathbb{A}_{ij}^{-1} \sum_k [\mathbb{A} - \mathbb{B}]_{jk} > 0$ .

<sup>36</sup> The reciprocal implication is shown merely for the sake of completeness.

(v) If there is a walk in  $\text{graph}(\mathbb{A} - \mathbb{B})$ ,  $x_i = y_0 \rightarrow \dots \rightarrow y_n \rightarrow x_j \rightarrow x_h$ , with  $y_m \in \bar{I}$  for all  $m$ ,  $x_j \in -I$  and  $h \neq i, j$ , then  $x_i \rightarrow x_h$  is an edge in  $\text{graph}(\text{Id} - \mathbb{A}^{-1}\mathbb{B})$ :

Using that  $\mathbb{A} - \mathbb{B} \leq \mathbb{A}$ , with equality for rows in  $\bar{I}$ , and  $\mathbb{A}_{ij}^{-1}[\mathbb{A} - \mathbb{B}]_{jh} < 0$  by assumption and (i), we have  $[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]_{ih} = [\mathbb{A}^{-1}(\mathbb{A} - \mathbb{B})]_{ih} = \sum_{k \neq j} \mathbb{A}_{ik}^{-1}[\mathbb{A} - \mathbb{B}]_{kh} + \mathbb{A}_{ij}^{-1}[\mathbb{A} - \mathbb{B}]_{jh} < \sum_{k \neq j} \mathbb{A}_{ik}^{-1}\mathbb{A}_{kh} = [\mathbb{A}^{-1}\mathbb{A}]_{ih} = 0$ , where we have used that  $\mathbb{A}_{jh} = \text{Id}_{jh} = 0$  for  $x_j \in -I$ ,  $j \neq h$ .

We turn now to the original claim. Let  $x_i \in J[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]^c$  such that there is a walk  $\omega$  from  $x_i$  to  $J[\mathbb{A} - \mathbb{B}]$  in  $\text{graph}(\mathbb{A} - \mathbb{B})$ . Assume without loss of generality that  $\omega$  has no closed subwalks. We want to find a shorter walk to  $J[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]$  in  $\text{graph}(\text{Id} - \mathbb{A}^{-1}\mathbb{B})$ .

Due to (ii), it must be either  $x_i \in \bar{I}$  or  $x_i \in -I$ , since states in  $\bar{I}^c \cap (-I)^c$  correspond to SDD rows. In the first case, (iv) implies that  $\omega$  must have the form  $\omega : x_i = y_0 \rightarrow \dots \rightarrow y_n \rightarrow x_j \rightarrow x_h \rightarrow \dots$  for some  $x_j \in (-I) \cap J[\mathbb{A} - \mathbb{B}]^c$ ,  $y_m \in \bar{I}$  for all  $m$  and  $h \neq j, i$  (recall that  $\omega$  has no closed subwalks). Then (v) tells us that one can skip states and walk from  $x_i$  to  $x_h$  with a single step, through  $\text{graph}(\text{Id} - \mathbb{A}^{-1}\mathbb{B})$ . Afterwards, one can continue with the same path for states in  $-I$  (if any), in light of (iii). The latter subwalk will either reach  $J[\mathbb{A} - \mathbb{B}]$  (hence,  $J[\text{Id} - \mathbb{A}^{-1}\mathbb{B}]$  as per (ii)) or go back to  $\bar{I}$  and repeat the same procedure. Since there are finitely many states,  $J[\mathbb{A} - \mathbb{B}]$  must eventually be reached in this way.

The case of  $x_i \in -I$  follows with the same reasoning. □

## References

1. Aïd, R., Basei, M., Callegaro, G., Campi, L., Vargiolu, T.: Nonzero-sum stochastic differential games with impulse controls: a verification theorem with applications. *Math. Oper. Res.* **45**(1), 205–232 (2019)
2. Azimzadeh, P., Bayraktar, E., Labahn, G.: Convergence of implicit schemes for Hamilton-Jacobi-Bellman quasi-variational inequalities. *SIAM J. Control Optim.* **56**(6), 3994–4016 (2018)
3. Aïd, R., Bernal, F., Mnif, M., Zabaljauregui, D., Zubelli, J.P.: A policy iteration algorithm for nonzero-sum stochastic impulse games. *ESAIM Proc. Surv.* **65**, 27–45 (2019)
4. Azimzadeh, P., Forsyth, P.A.: Weakly chained matrices, policy iteration, and impulse control. *SIAM J. Numer. Anal.* **54**(3), 1341–1364 (2016)
5. Altarovici, A., Reppen, M., Soner, H.M.: Optimal consumption and investment with fixed and proportional transaction costs. *SIAM J. Control Optim.* **55**(3), 1673–1710 (2017)
6. Azimzadeh, P.: Impulse control in finance: numerical methods and viscosity solutions, Ph.D. thesis, University of Waterloo, 2017
7. Azimzadeh, P.: A fast and stable test to check if a weakly diagonally dominant matrix is a nonsingular m-matrix. *Math. Comput.* **88**(316), 783–800 (2019)
8. Azimzadeh, P.: A zero-sum stochastic differential game with impulses, precommitment, and unrestricted cost functions. *Appl. Math. Optim.* **79**(2), 483–514 (2019)
9. Basei, M.: Optimal price management in retail energy markets: an impulse control problem with asymptotic estimates. *Math. Methods Oper. Res.* **89**(3), 355–383 (2019)
10. Basei, M., Cao, H., Guo, X.: Nonzero-sum stochastic games with impulse controls, arXiv preprint [arXiv:1901.08085](https://arxiv.org/abs/1901.08085) (2019)
11. Bokanowski, O., Maroso, S., Zidani, H.: Some convergence results for Howard’s algorithm. *SIAM J. Numer. Anal.* **47**(4), 3001–3026 (2009)
12. Barles, G., Souganidis, P.E.: Convergence of approximation schemes for fully nonlinear second order equations. *Asymptot. Anal.* **4**(3), 271–283 (1991)

13. Cadenillas, A., Choulli, T., Taksar, M., Zhang, L.: Classical and impulse stochastic control for the optimization of the dividend and risk policies of an insurance firm. *Math Financ.* **16**(1), 181–202 (2006)
14. Clift, S.: Linear and non-linear monotone methods for valuing financial options under two-factor, jump-diffusion models, Ph.D. thesis, (2007)
15. Chancelier, J.-P., Messaoud, M., Sulem, A.: A policy iteration algorithm for fixed point problems with nonexpansive operators. *Math. Methods Oper. Res.* **65**(2), 239–259 (2007)
16. Chancelier, J.-P., Øksendal, B., Sulem, A.: 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)
17. Cosso, A.: Stochastic differential games involving impulse controls and double-obstacle quasi-variational inequalities. *SIAM J. Control Optim.* **51**(3), 2102–2131 (2013)
18. Chang, D., Wang, H., Wu, Z.: Maximum principle for non-zero sum differential games of bsdes involving impulse controls. *Proceedings of the 32nd Chinese Control Conference, IEEE*, (2013) pp. 1564–1569
19. Cadenillas, A., Zapatero, F.: Optimal central bank intervention in the foreign exchange market. *J. Econ. Theory* **87**(1), 218–242 (1999)
20. De Angelis, T., Ferrari, G., Moriarty, J.: Nash equilibria of threshold type for two-player nonzero-sum games of stopping. *Ann. Appl. Probab.* **28**(1), 112–147 (2018)
21. El Asri, B., Mazid, S.: Zero-sum stochastic differential game in finite horizon involving impulse controls. *Appl. Math. Optim.* 1–33 (2018)
22. Egami, M.: A direct solution method for stochastic impulse control problems of one-dimensional diffusions. *SIAM J. Control Optim.* **47**(3), 1191–1218 (2008)
23. Eastham, J.F., Hastings, K.J.: Optimal impulse control of portfolios. *Math. Oper. Res.* **13**(4), 588–605 (1988)
24. Ferrari, G., Koch, T.: On a strategic model of pollution control. *Ann. Oper. Res.* **275**(2), 297–319 (2019)
25. Forsyth, P.A., Labahn, G.: Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *J. Comput. Financ.* **11**(2), 1 (2007)
26. Huang, Y., Forsyth, P.A., Labahn, G.: Combined fixed point and policy iteration for Hamilton-Jacobi-Bellman equations in finance. *SIAM J. Numer. Anal.* **50**(4), 1861–1882 (2012)
27. Huang, Y., Forsyth, P.A., Labahn, G.: Inexact arithmetic considerations for direct control and penalty methods: American options under jump diffusion. *Appl. Numer. Math.* **72**, 33–51 (2013)
28. Highman, N.J.: Accuracy and stability of numerical algorithms. SIAM, Philadelphia (2002)
29. Harsanyi, J.C., Selten, R.: A General Theory of Equilibrium Selection in Games, vol. 1. The MIT Press, Cambridge (1988)
30. Jeanblanc-Picqué, M.: Impulse control method and exchange rate. *Math. Financ.* **3**(2), 161–177 (1993)
31. Korn, R.: Some applications of impulse control in mathematical finance. *Math. Methods Oper. Res.* **50**(3), 493–518 (1999)
32. Mundaca, G., Øksendal, B.: Optimal stochastic intervention control with application to the exchange rate. *J. Math. Econ.* **2**(29), 225–243 (1998)
33. Øksendal, B., Sulem, A.: *Applied Stochastic Control of Jump Processes*, 3rd edn. Springer, Berlin (2009)
34. Reisinger, C., Zhang, Y.: A penalty scheme for monotone systems with interconnected obstacles: convergence and error estimates. *SIAM J. Numer. Anal.* **57**(4), 1625–1648 (2019)
35. Reisinger, C., Zhang, Y.: Error estimates of penalty schemes for quasi-variational inequalities arising from impulse control problems. *SIAM J. Control Optim.* **58**(1), 243–276 (2020)
36. Shivakumar, P.N., Chew, K.H.: A sufficient condition for nonvanishing of determinants. *Proceedings of the American mathematical society* 63–66 (1974)
37. Seydel, R.C.: Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions. *Stoch. Process. Appl.* **119**(10), 3719–3748 (2009)
38. Wang, J., Forsyth, P.A.: Maximal use of central differencing for Hamilton-Jacobi-Bellman PDEs in finance. *SIAM J. Numer. Anal.* **46**(3), 1580–1601 (2008)

39. Wang, S., Wu, Z.: Stochastic maximum principle for optimal control problems of forward-backward delay systems involving impulse controls. *J. Syst. Sci. Complex.* **30**(2), 280–306 (2017)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.