A hot-potato game under transient price impact

Alexander Schied* Tao Zhang*
Department of Mathematics
University of Mannheim
68131 Mannheim, Germany

First version: April 25, 2013
This version: September 2015

Abstract

We consider a Nash equilibrium between two high-frequency traders in a simple market impact model with transient price impact and additional quadratic transaction costs. Extending a result by Schöneborn (2008), we prove existence and uniqueness of the Nash equilibrium and show that for small transaction costs the high-frequency traders engage in a "hot-potato game", in which the same asset position is sold back and forth. We then identify a critical value for the size of the transaction costs above which all oscillations disappear and strategies become buy-only or sell-only. Numerical simulations show that for both traders the expected costs can be lower with transaction costs than without. Moreover, the costs can increase with the trading frequency when there are no transaction costs, but decrease with the trading frequency when transaction costs are sufficiently high. We argue that these effects occur due to the need of protection against predatory trading in the regime of low transaction costs.

Keywords: Hot-potato game, high-frequency trading, Nash equilibrium, transient price impact, market impact, predatory trading, M-matrix, inverse-positive matrix

1 Introduction

According to the Report [10] by CFTC and SEC on the Flash Crash of May 6, 2010, the events that lead to the Flash Crash included a large sell order of E-Mini S&P 500 contracts:

> ...a large Fundamental Seller (...) initiated a program to sell a total of 75,000 E-Mini contracts (valued at approximately $4.1 billion). ...[On another] occasion it took more than 5 hours for this large trader to execute the first 75,000 contracts of a large sell program. However, on May 6, when markets were already under stress, the Sell Algorithm chosen by the large Fundamental Seller to only target trading volume, and not price nor time, executed the sell program extremely rapidly in just 20 minutes.

*The authors acknowledge support by Deutsche Forschungsgemeinschaft through Research Grant SCHI 500/3-1

The report [10] furthermore suggests that a "hot-potato game" between high-frequency traders (HFTs) created artificial trading volume that at least contributed to the acceleration of the Fundamental Seller's trading algorithm:

> ... HFTs began to quickly buy and then resell contracts to each other—generating a "hot-potato" volume effect as the same positions were rapidly passed back and forth. Between 2:45:13 and 2:45:27, HFTs traded over 27,000 contracts, which accounted for about 49 percent of the total trading volume, while buying only about 200 additional contracts net.

See also Kirilenko, Kyle, Samadi, and Tuzun [15] and Easley, López da Prado, and O'Hara [11] for additional background.

Schöneborn [21] observed that the equilibrium strategies of two competing economic agents, who trade sufficiently fast in a simple market impact model with exponential decay of price impact, can exhibit strong oscillations. These oscillations have a striking similarity with the "hot-potato game" mentioned in [10] and [15]. In each trading period, one agent sells a large asset position to the other agent and buys a similar position back in the next period. The intuitive reason for this hot-potato game is to protect against possible *predatory trading* by the other agent. Here, predatory trading refers to the exploitation of the drift generated by the price impact of another agent. For instance, if the other agent is selling assets over a certain time interval, predatory trading would consist in shortening the asset at the beginning of the time interval and buying back when prices have depreciated through the sale of the other agent. Such strategies are "predatory" in the sense that their price impact decreases the revenues of the other agent and thus generate profit at the other agent's expense.

In this paper, we continue the investigation of the "hot-potato game". Our first contribution is to extend the result of Schöneborn [21] by identifying a unique Nash equilibrium for two competing agents within a larger class of adaptive trading strategies, for general decay kernels, and by giving an explicit formula for the equilibrium strategies. This explicit formula will be the starting point for our further mathematical and numerical analysis of the Nash equilibrium. Another new feature of our approach is the addition of quadratic transaction costs, which can be thought of temporary price impact in the sense of [6, 4] or as a transaction tax. The main goal of our paper is to study the impact of these additional transaction costs on equilibrium strategies. Theorem 2.8, our main result, precisely identifies a critical threshold θ^* for the size θ of these transaction costs at which all oscillations disappear. That is, for transactions $\theta \geq \theta^*$ certain "fundamental" equilibrium strategies consist exclusive of all buy trades or of all sell trades. For $\theta < \theta^*$, the "fundamental" equilibrium strategies will contain both buy and sell trades when the decay of price impact in between two trades is sufficiently small.

In addition, numerical simulations will exhibit some rather striking properties of equilibrium strategies. They reveal, for instance, that the expected costs of both agents can be a *decreasing* function of $\theta \in [0, \theta_0]$ when trading speed is sufficiently high. As a result, both agents can carry out their respective trades at a *lower cost* when there are transaction costs, compared to the situation without transaction costs. Even more interesting is the behavior of the costs as a function of the trading frequency. We will see that, for $\theta = \theta^*$, a higher trading speed can *decrease* expected trading costs, whereas the costs typically *increase* for sufficiently small θ. In particular the latter effect is surprising, because at first glance a higher trading frequency suggests that one has greater flexibility in the choice of a strategy and hence can become more cost efficient. So why are the costs then increasing in N? We will argue that the intuitive reason for this effect is that a higher trading frequency results in greater possibilities for predatory trading by the competitor and thus requires taking additional measures of protection against predatory trading.

This paper builds on several research developments in the existing literature. First, there are several papers on predatory trading such as Brunnermeier and Pedersen [8], Carlin et al. [9], Schöneborn and Schied [22], and the authors [20] dealing with Nash equilibria for several agents that are active in a market model with temporary and permanent price impact. A discrete-time market impact game with asymmetric information was analyzed by Moallemi et al. [17]. In contrast to these previous studies, the transient price impact model we use here goes back to Bouchaud et al. [7] and Obizhaeva and Wang [18]. It was further developed in [1, 2, 12, 3, 19], to mention only a few related papers. As first observed in [21], the qualitative features of Nash equilibria for transient price impact differ dramatically from those obtained in [9, 22, 20] for an Almgren–Chriss setting. We refer to [13, 16] for recent surveys on the price impact literature and extended bibliographies.

The paper is organized as follows. In Section 2.1 we explain our modeling framework. The existence and uniqueness theorem for Nash equilibria is stated in Section 2.2. In Section 2.3 we analyze the oscillatory behavior of equilibrium strategies. Here we will also state our main result, Theorem 2.8, on the critical threshold for the disappearance of oscillations. Our numerical results and their interpretation are presented in Section 2.3 and Section 2.4. Particularly, Section 2.4 contains the simulations for the behavior of the costs as a function of transaction costs and of trading frequency in the cases with and without transaction costs. In Section 2.5 we will analyze the asymptotic behavior of certain equilibrium strategies when the trading frequency tends to infinity. The proofs of our results are given in Section 3. We conclude in Section 4.

2 Statement of results

2.1 Modeling framework

We consider two financial agents, X and Y, who are active in a market impact model for one risky asset. Market impact will be transient and modeled as in [3]; see also [7, 18, 2, 12, 19] for closely related or earlier versions of this model, which is sometimes called a propagator model. When none of the two agents is active, asset prices are described by a right-continuous martingale[1] $S^0 = (S^0_t)_{t\geq 0}$ on a filtered probability space $(\Omega, (\mathscr{F}_t)_{t\geq 0}, \mathscr{F}, \mathbb{P})$, for which $\mathscr{F}_0$ is $\mathbb{P}$-trivial. The process S^0 is often called the unaffected price process. Trading takes place at the discrete trading times of a *time grid* $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$, where $0 = t_0 < t_1 < \cdots < t_N = T$. Both agents are assumed to use trading strategies that are admissible in the following sense.

Definition 2.1. Suppose that a time grid $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$ is given. An *admissible trading strategy* for $\mathbb{T}$ and $Z_0 \in \mathbb{R}$ is a vector $\boldsymbol{\zeta} = (\zeta_0, \dots, \zeta_N)$ of random variables such that

(a) each ζ_i is $\mathscr{F}_{t_i}$-measurable and bounded, and

(b) $Z_0 = \zeta_0 + \cdots + \zeta_N$ $\mathbb{P}$-a.s.

[1]The martingale assumption is natural from an economic point of view, because we are interested here in high-frequency trading over short time intervals $[0, T]$. See also the discussion in [3] for additional arguments.

The set of all admissible strategies for given $\mathbb{T}$ and Z_0 is denoted by $\mathscr{X}(Z_0, \mathbb{T})$.

For $\boldsymbol{\zeta} \in \mathscr{X}(Z_0, \mathbb{T})$, the value of ζ_i is taken as the number of shares traded at time t_i, with a positive sign indicating a sell order and a negative sign indicating a purchase. Thus, the requirement (b) in the preceding definition can be interpreted by saying that Z_0 is the inventory of the agent at time $0 = t_0$ and that by time $t_N = T$ (e.g., the end of the trading day) the agent must have a zero inventory. The assumption that each ζ_i is bounded can be made without loss of generality from an economic point of view.

When the two agents X and Y apply respective strategies $\boldsymbol{\xi} \in \mathscr{X}(X_0, \mathbb{T})$ and $\boldsymbol{\eta} \in \mathscr{X}(Y_0, \mathbb{T})$, the asset price is given by

$$S_t^{\boldsymbol{\xi},\boldsymbol{\eta}} = S_t^0 - \sum_{t_k < t} G(t - t_k)(\xi_k + \eta_k), \tag{1}$$

where $G : \mathbb{R}_+ \to \mathbb{R}_+$ is a function called the *decay kernel*. Thus, at each time $t_k \in \mathbb{T}$, the combined trading activities of the two agents move the current price by the amount $-G(0)(\xi_k + \eta_k)$. At a later time $t > t_k$, this price impact will have changed to $-G(t - t_k)(\xi_k + \eta_k)$. From an economic point of view it would be reasonable to assume that G is nonincreasing, but this assumption is not essential for our results to hold mathematically. But we do assume throughout this paper that the function $t \mapsto G(|t|)$ is strictly positive definite in the sense of Bochner: For all $n \in \mathbb{N}$, $t_1, \dots, t_n \in \mathbb{R}$, and $x_1, \dots, x_n \in \mathbb{R}$ we have

$$\sum_{i,j=1}^{n} x_i x_j G(|t_i - t_j|) \geq 0, \quad \text{with equality if and only if } x_1 = \cdots = x_n = 0. \tag{2}$$

As observed in [3], this assumption rules out the existence of price manipulation strategies in the sense of Huberman and Stanzl [14]. It is satisfied as soon as G is convex, nonincreasing, and nonconstant; see, e.g., [3, Proposition 2] for a proof.

Let us now discuss the definition of the liquidation costs incurred by each agent. When only one agent, say X, places a nonzero order at time t_k, then we are in the situation of [3] and the price is moved linearly from $S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}$ to $S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} := S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}} - G(0)\xi_k$. The order ξ_k is therefore executed at the average price $\frac{1}{2}(S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} + S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}})$ and consequently incurs the following expenses:

$$-\frac{1}{2}\big(S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} + S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\big)\xi_k = \frac{G(0)}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k.$$

Suppose now that the order η_k of agent Y is executed immediately after the order ξ_k. Then the price is moved linearly from $S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}}$ to $S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} - G(0)\eta_k$, and the order of agent Y incurs the expenses

$$-\frac{1}{2}\big(S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} + S_{t_k+}^{\boldsymbol{\xi},\boldsymbol{\eta}} - G(0)\eta_k\big)\eta_k = \frac{G(0)}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + G(0)\xi_k\eta_k.$$

So greater latency results in the additional cost term $G(0)\xi_k\eta_k$ for agent Y. Clearly, this term would appear in the expenses of agent X when the roles of X and Y are reversed. In the sequel, we are going to assume that none of the two agents has an advantage in latency over the other. Therefore, if both agents place nonzero orders at time t_k, execution priority is given to that agent who wins an independent coin toss.

In addition to the liquidation costs motivated above, we will also impose that each trade ζ_k incurs quadratic transaction costs of the form $\theta\zeta_k^2$, where θ is a nonnegative parameter.

Definition 2.2. Suppose that $\mathbb{T} = \{t_0, t_1, \dots, t_N\}$, X_0 and Y_0 are given. Let furthermore $(\varepsilon_i)_{i=0,1,\dots}$ be an i.i.d. sequence of Bernoulli $(\frac{1}{2})$-distributed random variables that are independent of $\sigma(\bigcup_{t\geq 0}\mathscr{F}_t)$. Then the *costs of* $\boldsymbol{\xi} \in \mathscr{X}(X_0, \mathbb{T})$ *given* $\boldsymbol{\eta} \in \mathscr{X}(Y_0, \mathbb{T})$ are defined as

$$\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}) = X_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{G(0)}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k G(0)\xi_k\eta_k + \theta\xi_k^2 \Big) \tag{3}$$

and the *costs of* $\boldsymbol{\eta}$ *given* $\boldsymbol{\xi}$ are

$$\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}) = Y_0 S_0^0 + \sum_{k=0}^{N} \Big(\frac{G(0)}{2}\eta_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\eta_k + (1-\varepsilon_k) G(0)\xi_k\eta_k + \theta\eta_k^2 \Big).$$

The term $X_0 S_0^0$ corresponds to the book value of the position X_0 at time $t = 0$. If the position X_0 could be liquidated at book value, one would incur the expenses $-X_0 S_0^0$. Therefore, the liquidation costs as defined in (3) are the difference of the actual accumulated expenses, as represented by the sum on the right-hand side of (3), and the expenses for liquidation at book value. The following two remarks provide further comments on our modeling assumptions.

Remark 2.3. The market impact model we are using here has often been linked to the placement of market orders in a block-shaped limit order book, and a bid-ask spread is sometimes added to the model so as to make this interpretation more feasible [18, 1]. For a strategy consisting exclusively of market orders, the bid-ask spread will lead to an additional fee that should be reflected in the corresponding cost functional. In reality, however, most strategies will involve a variety of different order types and one should think of the costs (3) as the costs *averaged* over order types, as is often done in the market impact literature. For instance, while one may have to pay the spread when placing a market order, one essentially earns it back when a limit order is executed. Moreover, high-frequency traders often have access to a variety of more exotic order types, some of which can pay rebates when executed. It is also possible to use crossing networks or dark pools in which orders are executed at mid price. So, for a setup of high-frequency trading, taking the bid-ask spread as zero in (1) is probably more realistic than modeling every single order as a market order and to impose the fees. The existence of hot-potato games in real-world markets, such as the one quoted from [10] in the Introduction, can be regarded as an empirical justification of the zero-spread assumption, because such a trading behavior could never be profitable if each trader had to pay the full spread upon each execution of an order.

Remark 2.4. We admit that we have chosen quadratic transaction costs because this choice makes our model mathematically tractable. Yet, there are several aspects why quadratic transaction costs may not be completely implausible from an economic point of view. For instance, these costs can be regarded as arising from temporary price impact in the spirit of [6, 4], which is also quadratic in order size. Moreover, these costs can model a transaction tax that is subject to tax progression. With such a tax, small orders, such as those placed by small investors, are taxed at a lower rate than large orders, which may be placed with the intention of moving the market. Finally, quadratic transaction costs differ from proportional transaction costs in a qualitative manner only when order size tends to zero. As long as the sizes of nonzero orders are bounded away from zero, as it must necessarily be the case in a model with finitely many trading dates, quadratic transaction costs can be replaced by a transaction cost function that is linear around the origin without changing the actual costs of a strategy. It is therefore reasonable to expect a similar qualitative behavior also for proportional transaction costs.

2.2 Nash equilibrium

We now consider agents who need to liquidate their current inventory within a given time frame and who are aiming to minimize the expected costs over admissible strategies. The need for liquidation can arise due to various reasons. For instance, Easley, López da Prado, and O'Hara [11] argue that the toxicity of the order flow preceding the Flash Crash of May 6, 2010, has led the inventory of several high-frequency market makers to grow beyond their risk limits, thus forcing them to unload this inventory.

When just a single agent is considered, the minimization of the expected execution costs is a well-studied problem; we refer to [3] for an analysis within our current modeling framework. Here we are going to investigate the optimal strategies of our two agents, X and Y, under the assumption that both have full knowledge of the other's strategy and maximize the expected costs of their strategies accordingly. In this situation, it is natural to define optimality through the following notion of a Nash equilibrium.

Definition 2.5. For given time grid $\mathbb{T}$ and initial values $X_0, Y_0 \in \mathbb{R}$, a *Nash equilibrium* is a pair $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*)$ of strategies in $\mathscr{X}(X_0,\mathbb{T}) \times \mathscr{X}(Y_0,\mathbb{T})$ such that

$$\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \inf_{\boldsymbol{\xi}\in\mathscr{X}(X_0,\mathbb{T})} \mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,] \qquad \text{and} \qquad \mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,] = \inf_{\boldsymbol{\eta}\in\mathscr{X}(Y_0,\mathbb{T})} \mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,].$$

To state our formula for this Nash equilibrium, we need to introduce the following notation. For a fixed time grid $\mathbb{T} = \{t_0, \dots, t_N\}$, we define the $(N+1)\times(N+1)$-matrix Γ by

$$\Gamma_{i,j} = G(|t_{i-1} - t_{j-1}|), \qquad i, j = 1, \dots, N+1, \tag{4}$$

and for $\theta \geq 0$ we introduce

$$\Gamma_\theta := \Gamma + 2\theta\,\mathrm{Id}. \tag{5}$$

We furthermore define the lower triangular matrix $\widetilde{\Gamma}$ by

$$\widetilde{\Gamma}_{ij} = \begin{cases} \Gamma_{ij} & \text{if } i > j, \\ \frac{1}{2}G(0) & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

We will write $\mathbf{1}$ for the vector $(1, \dots, 1)^\top \in \mathbb{R}^{N+1}$. A strategy $\boldsymbol{\zeta} = (\zeta_0, \dots, \zeta_N) \in \mathscr{X}(Z_0, \mathbb{T})$ will be identified with the $(N+1)$-dimensional random vector $(\zeta_0, \dots, \zeta_N)^\top$. Conversely, any vector $\boldsymbol{z} = (z_1, \dots, z_{N+1})^\top \in \mathbb{R}^{N+1}$ can be identified with the deterministic strategy $\boldsymbol{\zeta}$ with $\zeta_k = z_{k+1}$. We also define the two vectors

$$\begin{aligned} \boldsymbol{v} &= \frac{1}{\mathbf{1}^\top(\Gamma_\theta + \widetilde{\Gamma})^{-1}\mathbf{1}}(\Gamma_\theta + \widetilde{\Gamma})^{-1}\mathbf{1} \\ \boldsymbol{w} &= \frac{1}{\mathbf{1}^\top(\Gamma_\theta - \widetilde{\Gamma})^{-1}\mathbf{1}}(\Gamma_\theta - \widetilde{\Gamma})^{-1}\mathbf{1}. \end{aligned} \tag{7}$$

It will be shown in Lemma 3.2 below that the matrices $\Gamma_\theta + \widetilde{\Gamma}$ and $\Gamma_\theta - \widetilde{\Gamma}$ are indeed invertible and that the denominators in (7) are strictly positive under our assumption (2) that $G(|\cdot|)$ is strictly positive definite. Recall that we assume (2) throughout this paper.

In the case $G(t) = \gamma + \lambda e^{-\rho t}$ for constants $\gamma \geq 0$ and $\lambda, \rho > 0$, the existence of a unique Nash equilibrium in the class of deterministic strategies was established in Theorem 9.1 of [21]. Our subsequent Theorem 2.6 extends this result in a number of ways: we allow for general positive definite decay kernels, include transaction costs, give an explicit form of the deterministic Nash equilibrium, and show that this Nash equilibrium is also the unique Nash equilibrium in the class of adapted strategies. Our explicit formula for the equilibrium strategies will be the starting point for our further mathematical and numerical analysis of the Nash equilibrium. Also our proof is different from the one in [21], which works only for the specific decay kernel $G(t) = \lambda e^{-\rho t} + \gamma$.

Theorem 2.6. *For any time grid $\mathbb{T}$ and initial values X_0, $Y_0 \in \mathbb{R}$, there exists a unique Nash equilibrium $(\boldsymbol{\xi}^*, \boldsymbol{\eta}^*) \in \mathscr{X}(X_0, \mathbb{T}) \times \mathscr{X}(Y_0, \mathbb{T})$. The optimal strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ are deterministic and given by*

$$\begin{aligned} \boldsymbol{\xi}^* &= \frac{1}{2}(X_0 + Y_0)\boldsymbol{v} + \frac{1}{2}(X_0 - Y_0)\boldsymbol{w}, \\ \boldsymbol{\eta}^* &= \frac{1}{2}(X_0 + Y_0)\boldsymbol{v} - \frac{1}{2}(X_0 - Y_0)\boldsymbol{w}. \end{aligned} \tag{8}$$

The formula (8) shows that the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ form a basis for all possible equilibrium strategies. It follows that in analyzing the Nash equilibrium it will be sufficient to study the two cases $\boldsymbol{\xi}^* = \boldsymbol{v} = \boldsymbol{\eta}^*$ for $X_0 = 1 = Y_0$ and $\boldsymbol{\xi}^* = \boldsymbol{w} = -\boldsymbol{\eta}^*$ for $X_0 = 1 = -Y_0$.

2.3 The hot-potato game

We now turn toward a qualitative analysis of the equilibrium strategies. Due to the computational complexity of this task, we will concentrate here on the case of an exponential decay kernel with additional permanent price impact:

$$G(t) = \lambda e^{-\rho t} + \gamma \qquad \text{for constants } \lambda, \rho > 0 \text{ and } \gamma \geq 0. \tag{9}$$

It is well known that this class of decay kernels satisfies our assumption (2) (see, e.g., [3, Example 1]).

By means of numerical simulations and the analysis of a particular example, Schöneborn [21, Section 9.3] observed that the equilibrium strategies may exhibit strong oscillations when $\theta = 0$. The subsequent proposition implies that for $\gamma = 0$ such oscillations will always occur in a Nash equilibrium with $X_0 = -Y_0$ if θ is sufficiently small and the trading frequency is sufficiently high. Note that this Nash equilibrium is completely determined by the vector $\boldsymbol{w}$. Throughout this and the following sections, we will concentrate on *equidistant time grids*,

$$\mathbb{T}_N := \left\{ \frac{kT}{N} \,\middle|\, k = 0, 1, \ldots, N \right\}, \qquad N \in \mathbb{N}. \tag{10}$$

Proposition 2.7. *Suppose that G is of the form* (9) *with $\gamma = 0$ and that $T > 0$ is fixed. Then there exists $N_0 \in \mathbb{N}$ such that for each $N \geq N_0$ there exists $\delta > 0$ so that for $0 \leq \theta < \delta$ the entries of the vector $\boldsymbol{w} = (w_1, \ldots, w_{N+1})$ are nonzero and have alternating signs.*

We refer to the right-hand panel of Figure 1 for an illustration of the oscillations of the vector $\boldsymbol{w}$. As shown in the left-hand panel of the same figure, similar oscillations occur for the vector

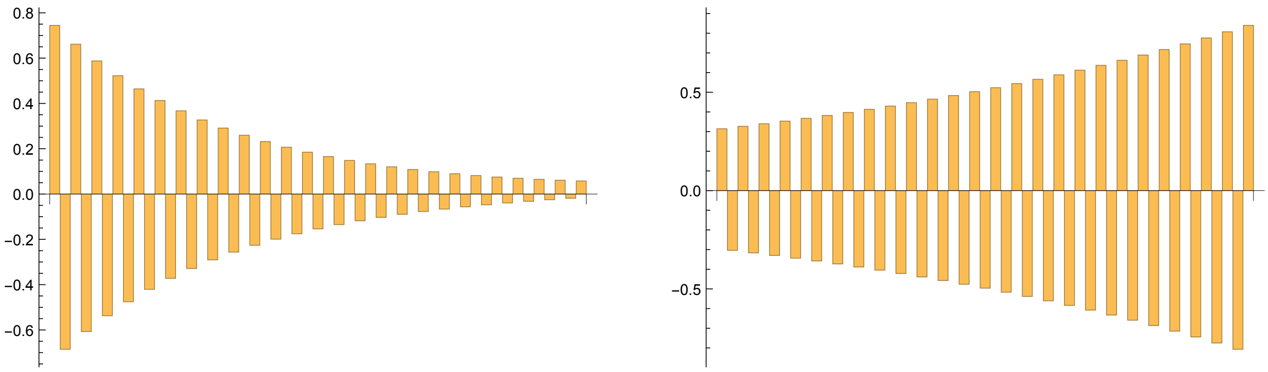

Figure 1: Vectors $\boldsymbol{v}$ (left) and $\boldsymbol{w}$ (right) for the equidistant time grid $\mathbb{T}_{50}$, $G(t) = e^{-t}$, $\theta = 0$, and $T = 1$. By (8), $(\boldsymbol{v}, \boldsymbol{v})$ is the equilibrium for $X_0 = Y_0 = 1$, and $(\boldsymbol{w}, -\boldsymbol{w})$ is the equilibrium for $X_0 = -Y_0 = 1$. Yet, some individual components of both $\boldsymbol{v}$ and $\boldsymbol{w}$ exceed in either direction 60% of the sizes of the initial positions X_0 and Y_0.

$\boldsymbol{v}$ and hence for equilibria with arbitrary initial conditions. The mathematical analysis for $\boldsymbol{v}$, however, is much harder than for $\boldsymbol{w}$, and at this time we are not able to prove a result that could be an analogue of Proposition 2.7 for the vector $\boldsymbol{v}$. The existence of oscillations of $\boldsymbol{w}$ and $\boldsymbol{v}$ is also not limited to exponential decay kernels as can be seen from numerical experiments; see Figure 2 for an illustration. We refer to Remark 2.9 for a possible financial interpretation of the oscillations arising in the hot-potato game. Already here we point out that for a single financial agent optimal strategies will always be buy-only or sell-only for the decay kernels used in Figures 1 and 2 (see [3, Theorem 1]). Therefore, the oscillations in our two-agent setting that are observed in these figures must necessarily result from the interaction of both agents.

We can now turn to presenting the main mathematical result of this paper. It is concerned with the cease of oscillations of both $\boldsymbol{v}$ and $\boldsymbol{w}$ when the parameter θ increases. Intuitively it is clear that increased transaction costs will penalize oscillating strategies and thus lead to a smoothing of the equilibrium strategies. As a matter of fact, one can see in Figure 3 that for $\theta = 2$ all oscillations have disappeared so that equilibrium strategies are then buy-only or sell-only. One can therefore wonder whether between $\theta = 0$ and $\theta = 2$ there might be a *critical value* θ^* at which all oscillations of $\boldsymbol{v}$ and $\boldsymbol{w}$ disappear but below which oscillations are present. That is, for $\theta \geq \theta^*$ all equilibrium strategies should be either buy-only or sell-only, while for $\theta < \theta^*$ equilibrium strategies should contain both buy and sell trades (at least for certain values of N and ρ). The following theorem confirms that such a critical value θ^* does indeed exist. We can even determine its precise value.

Theorem 2.8. *Suppose that G is as in* (9) *and $\mathbb{T}_N$ denotes the equidistant time grid* (10)*. Then the following conditions are equivalent.*

(a) *For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{v}$ are nonnegative.*

(b) *For every $N \in \mathbb{N}$ and $\rho > 0$, all components of $\boldsymbol{w}$ are nonnegative.*

(c) $\theta \geq \theta^* = (\lambda + \gamma)/4$.

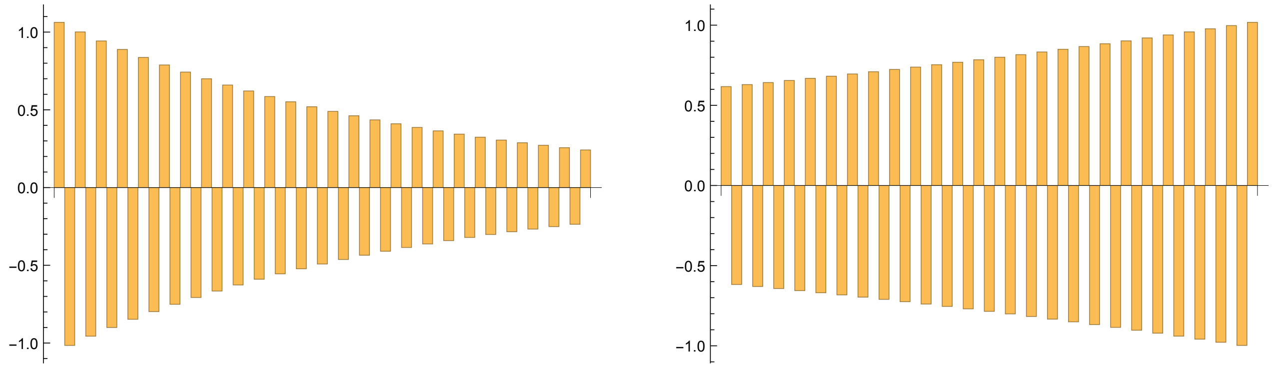

Figure 2: Vectors $\boldsymbol{v}$ (left) and $\boldsymbol{w}$ (right) for price impact decaying according to the power-law $G(t) = (1+t)^{-0.5}$, as suggested, e.g., in [12]. The remaining parameters are as in Figure 1. Note the qualitative similarity to the corresponding strategies for exponential decay of price impact in Figure 1.

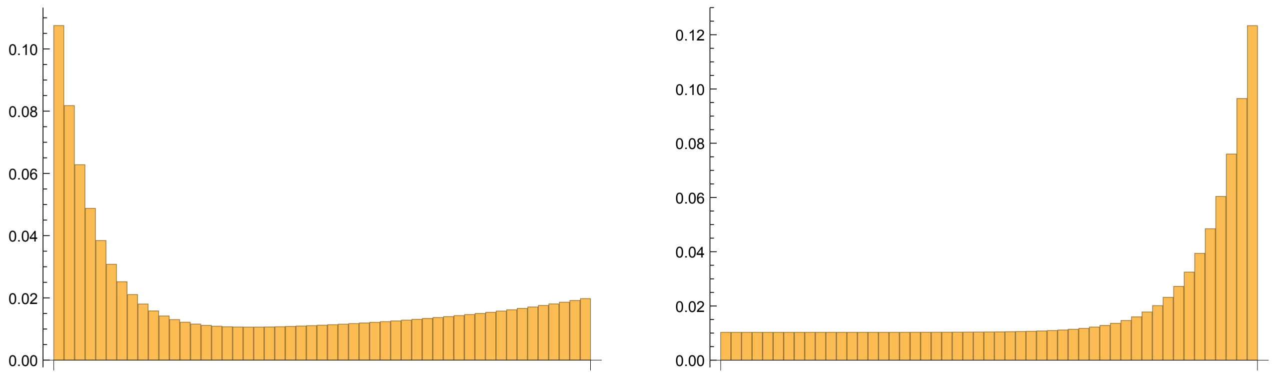

Figure 3: Vectors $\boldsymbol{v}$ (left) and $\boldsymbol{w}$ (right) for the equidistant time grid $\mathbb{T}_{50}$, $G(t) = e^{-t}$, $\theta = 2$, and $T = 1$.

Remark 2.9. In this remark we will discuss a possible financial explanation for the oscillations of equilibrium strategies observed for small values of θ. As mentioned above, the source for these oscillations must necessarily lie in the interaction between the two agents. As observed in previous studies on multi-agent equilibria in price impact models such as [8, 9, 22], the dominant form of interaction between two players is *predatory trading*, which consists in the exploitation of price impact generated by another agent. Such strategies are "predatory" in the sense that they generate profit by simultaneously decreasing the other agent's revenues. Since predators prey on the drift created by the price impact of a large trade, protection against predatory trading requires the erasion of previously created price impact. Under transient price impact, the price impact of an earlier trade, say ζ_0, can be erased by placing an order ζ_1 of the opposite side. For instance, taking $\zeta_1 := -\zeta_0 G(t_1 - t_0)$ will completely eliminate the price impact of ζ_0 while the combined trades execute a total of $\xi_0(1 - G(t_1 - t_0))$ shares. In this sense, oscillating strategies can be understood as a protection against predatory trading by opponents (see also [21, p.150]).

Remark 2.10. Alfonsi et al. [3] discovered oscillations for the trade execution strategies of a *single* trader under transient price impact when price impact does not decay as a convex function of time. These oscillations, however, result from an attempt to exploit the delay in market response to a large trade, and they disappear when price impact decays as a convex function of time [3, Theorem 1]. In particular, when there is just one agent active and G is convex, nonincreasing, and nonconstant (which is, e.g., the case under assumption (9)), then for each time grid $\mathbb{T}$ there exists a unique optimal strategy, which is either buy-only or sell-only. When (9) holds and $\theta = 0$, this strategy is known explicitly; see [1].

2.4 The impact of transaction costs and trading frequency on the expected costs

Due to our explicit formulas (7) and (8), it is easy to analyze the Nash equilibrium numerically. These numerical simulations exhibit several striking effects in regards to monotonicity properties of the expected costs.

In Figure 4 we have plotted the expected costs $\mathbb{E}[\mathscr{C}_{\mathbb{T}_N}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)] = \mathbb{E}[\mathscr{C}_{\mathbb{T}_N}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)]$ for $X_0 = Y_0$, $G(t) = e^{-t}$, and $T = 1$ as a function of the trading frequency, N. The first observation one probably makes when looking at this plot is the fact that for $\theta = 0$ the expected costs exhibit a sawtooth-like pattern; they alternate between two increasing trajectories, depending on whether N is odd or even. These alternations are due to the oscillations of the optimal strategies, which also alternate with N. As can be seen from the figure, the sawtooth pattern disappears already for very small values of θ such as for $\theta = 0.08$.

A more interesting observation is the fact that for $\theta = 0$ and $\theta = 0.08$ the expected costs $\mathbb{E}[\mathscr{C}_{\mathbb{T}_{2N}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)]$ (or alternatively $\mathbb{E}[\mathscr{C}_{\mathbb{T}_{2N+1}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)]$) are *increasing* in N. This fact is surprising because a higher trading frequency should normally lead to a larger class of admissible strategies. As a result, traders have greater flexibility in choosing a strategy and in turn should be able to pick more cost efficient strategies. So why are the costs then increasing in N? The intuitive explanation is that a higher trading frequency increases also the possibility for the competitor to conduct predatory strategies at the expense of the other agent (see Remark 2.9). In reaction, this other agent needs to take stronger protective measures against predatory trading. As discussed in Remark 2.9, protection against predatory trading can be obtained by erasing (part of) the previously created price impact through placing an order of the opposite side. The result is an oscillatory strategy, whose expected costs increase with the number of its oscillations.

Still in Figure 4, the expected costs $\mathbb{E}[\mathscr{C}_{\mathbb{T}_N}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)]$ for the case $\theta = \theta^* = 0.25$ exhibit a very different behavior. They no longer alternate in N and are *decreasing* as a function of the trading frequency. The intuitive explanation is that transaction costs of size $\theta^* = 0.25$ discourage predatory trading to a large extend, so that agents can now benefit from a higher trading frequency and pick ever more cost-efficient strategies as N increases.

The most surprising observation in Figure 4 is the fact that for sufficiently large N the expected costs for $\theta = 0.08$ and for $\theta = \theta^* = 0.25$ fall below the expected costs for $\theta = 0$. That is, for sufficiently large trading frequency, *adding transaction costs can decrease the expected costs of all market participants* (recall that for $X_0 = Y_0$ both agents have the same optimal strategies and, hence, the same expected costs). This fact is further illustrated in Figure 5, which exhibits a very steep initial decrease of the expected costs as a function of θ. After a minimum of the expected costs is reached at $\theta \approx 0.06$, there is a slow and steady increase of the costs with an approximate slope of 0.002.

The key to understanding the behavior of expected equilibrium costs as a function of trading frequency and transaction costs rests in the interpretation of the oscillations in equilibrium strategies as a protection against predatory trading by the opponent (see Remark 2.9). Note that a predatory trading strategy is necessarily a "round trip", i.e., a strategy with zero inventory at $t = 0$ and $T = 0$ (the strategy of a predatory trader with nonzero initial position would consist of a superposition of a predatory round trip and a liquidation strategy for the initial position). It therefore must consist of a buy and a sell component and is hence stronger penalized by an increase in transaction costs than a buy-only or sell-only strategy. As a result, increasing transaction costs leads to an overall reduction of the proportion of predatory trades in equilibrium. In consequence, both agents in our model can reduce their protection against predatory trading and therefore use more efficient strategies to carry out their trades. They can thus fully benefit from higher trading frequencies, which leads to the observed decrease of expected costs as a function of N if θ is sufficiently large. Moreover, for appropriate values of $\theta > 0$, the benefit of increased efficiency outweighs the price to be paid in higher transaction costs and so an overall reduction of costs is achieved.

2.5 Analysis of the high-frequency limit

In the sequel, we analyze the possible convergence of the equilibrium strategies when the trading frequency tends to infinity. To this end, we consider the equidistant time grids $\mathbb{T}_N$ as defined in (10) for varying $N \in \mathbb{N}$ and write $\boldsymbol{v}^{(N)} = (v_1^{(N)}, \dots, v_{N+1}^{(N)})$ and $\boldsymbol{w}^{(N)} = (w_1^{(N)}, \dots, w_{N+1}^{(N)})$ for the vectors in (7) to make the dependence on N explicit. We start with the following proposition, which analyzes the convergence of the individual components of $\boldsymbol{w}^{(N)}$ when $N \uparrow \infty$. By (8), a Nash equilibrium with $X_0 = -Y_0$ is completely determined by $\boldsymbol{w}^{(N)}$.

Proposition 2.11. *Suppose that n is fixed.*

(a) *When $\theta = 0$, we have*

$$\lim_{N\uparrow\infty} w_n^{(2N)} = (-1)^{n+1} \frac{2a}{2\rho T + a + 1} \quad \textit{and} \quad \lim_{N\uparrow\infty} w_n^{(2N+1)} = (-1)^n \frac{2a}{2\rho T - a + 1}, \tag{11}$$

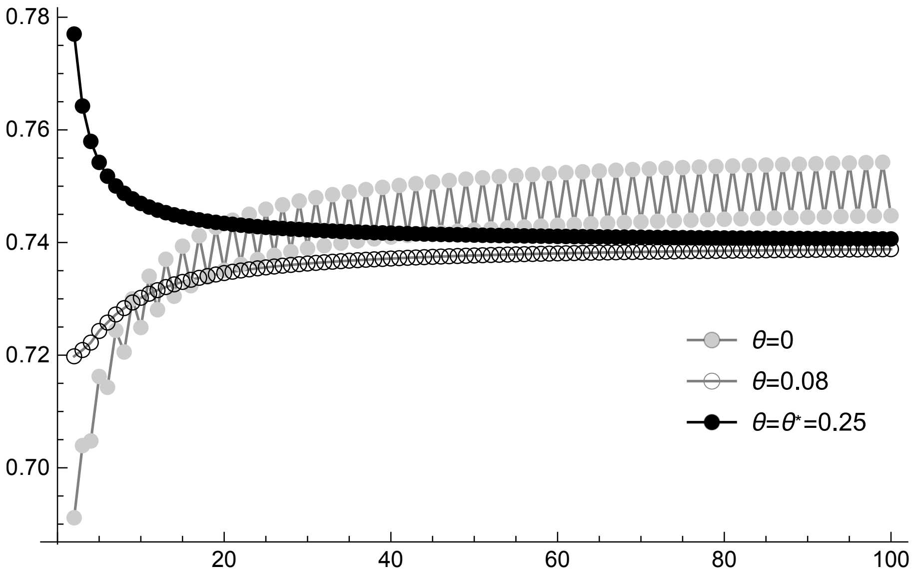

Figure 4: Expected costs $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}_N}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \mathbb{E}[\,\mathscr{C}_{\mathbb{T}_N}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,]$ for various values of θ as a function of trading frequency, N, with the equidistant time grid $\mathbb{T}_N$, $T = 1$, $G(t) = e^{-t}$, and $X_0 = Y_0 = 1$.

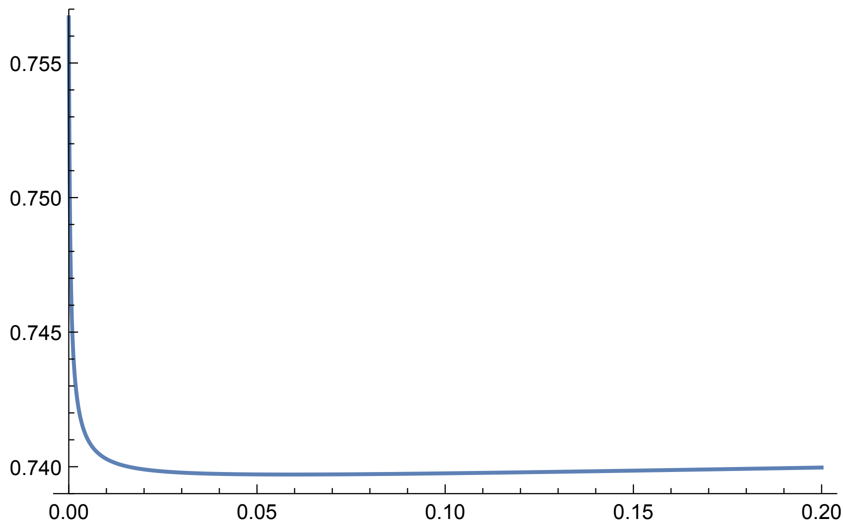

Figure 5: Expected costs $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}_{501}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,] = \mathbb{E}[\,\mathscr{C}_{\mathbb{T}_{501}}(\boldsymbol{\eta}^*|\boldsymbol{\xi}^*)\,]$ as a function of θ. The costs decrease steeply from the value 0.7567 at $\theta = 0$ until a minimum value of about 0.7397 at $\theta = 0.06$. From then on there is a moderate and almost linear increase with, e.g., a value of 0.7407 at $\theta = 0.5$. This increase corresponds to a slope of approximately 0.002. We took the equidistant time grid $\mathbb{T}_{501}$, initial values $X_0 = Y_0 = 1$, and $\lambda = \rho = 1$.

as well as

$$\lim_{N\uparrow\infty} w^{(2N)}_{2N+1-n} = (-1)^n \frac{2}{2\rho T + a + 1} \quad \textit{and} \quad \lim_{N\uparrow\infty} w^{(2N+1)}_{2N+2-n} = (-1)^n \frac{2}{2\rho T - a + 1}. \tag{12}$$

(b) *When $\theta > 0$, we have*

$$\lim_{N\uparrow\infty} w^{(N)}_n = 0,$$

and

$$\lim_{N\uparrow\infty} w^{(N)}_{N+1-n} = \Big(\frac{4\theta - \lambda}{4\theta + \lambda}\Big)^n \frac{2\lambda}{(\rho T + 1)(4\theta + \lambda)}. \tag{13}$$

The preceding proposition gives further background on the oscillations of equilibrium strategies in the regime $\theta < \theta^*$. In particular, (11) shows that in a Nash equilibrium with $X_0 = -Y_0$ and with $\theta = 0$ the trades of both agents asymptotically oscillate between $\pm const$ and that the sign of each trade also depends on whether N is odd or even. Moreover, (13) implies that, for $K \in \mathbb{N}$ fixed and $N \uparrow \infty$, the terminal K trades in an equilibrium strategy asymptotically oscillate between $\pm const$ if and only if $\theta < \lambda/4$.

Now we consider the Nash equilibrium $(\boldsymbol{\xi}^{*,(N)}, \boldsymbol{\eta}^{*,(N)})$ with initial positions X_0, Y_0 and time grid $\mathbb{T}_N$. We define the asset positions of the two agents via

$$X^{(N)}_t := X_0 - \sum_{k=1}^{\lceil \frac{Nt}{T} \rceil} \xi^{*,(N)}_k, \qquad Y^{(N)}_t := Y_0 - \sum_{k=1}^{\lceil \frac{Nt}{T} \rceil} \eta^{*,(N)}_k, \qquad t \geq 0. \tag{14}$$

In the case $X_0 = -Y_0$, we have $X^{(N)} = -Y^{(N)} = X_0 W^{(N)}$, where

$$W^{(N)}_t = 1 - \sum_{k=1}^{\lceil \frac{Nt}{T} \rceil} w^{(N)}_k, \qquad t \geq 0. \tag{15}$$

Proposition 2.12. *When $\theta > 0$, we have for $t < T$*

$$\lim_{N\uparrow\infty} W^{(N)}_t = \frac{\rho(T - t) + 1}{\rho T + 1} \tag{16}$$

and $W_t = 0$ for $t > T$.

Note that the limiting function in (16) is independent of θ as long as $\theta > 0$. Let

$$V^{(N)}_t = 1 - \sum_{k=1}^{\lceil \frac{Nt}{T} \rceil} \boldsymbol{v}^{(N)}_k, \qquad t \geq 0. \tag{17}$$

The asymptotic analysis for $V^{(N)}$ is much more difficult than for $W^{(N)}$ and, at this time, we are not able to prove any quantitative results. But the numerical simulation in Figure 6 suggests that $V^{(N)}$ converges for $\theta \geq \theta^*$ to a function $V^{(\infty)}$, which has a jump at $t = 0$ and is otherwise a nonlinear function of time.

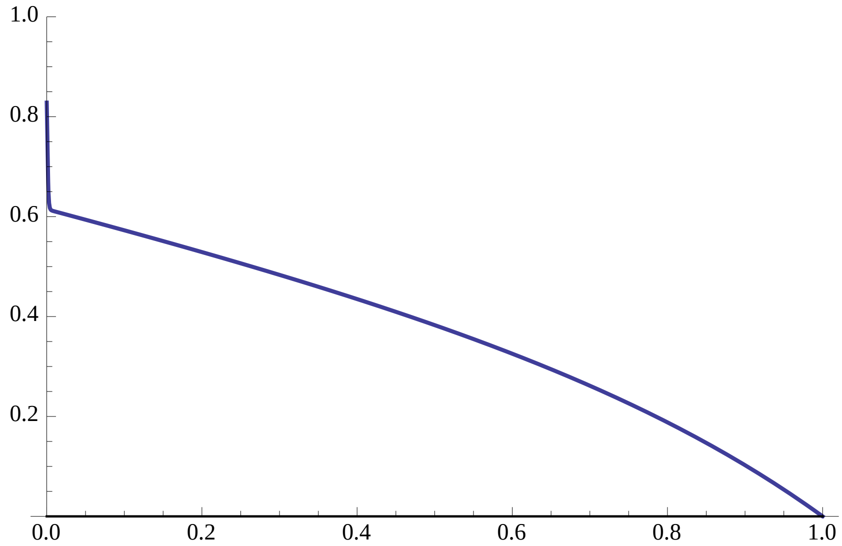

Figure 6: The function $V^{(N)}$ for $N = 1000$ with parameters $\theta = 0.5$ and $\lambda = \rho = 1$.

3 Proofs

3.1 Proof of Theorem 2.6

Lemma 3.1. *The expected costs of an admissible strategy* $\boldsymbol{\xi} \in \mathscr{X}(X_0, \mathbb{T})$ *given another admissible strategy* $\boldsymbol{\eta} \in \mathscr{X}(Y_0, \mathbb{T})$ *are*

$$\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,] \quad = \quad \mathbb{E}\Big[\,\frac{1}{2}\boldsymbol{\xi}^\top \Gamma_\theta \boldsymbol{\xi} + \boldsymbol{\xi}^\top \widetilde{\Gamma} \boldsymbol{\eta}\,\Big]. \tag{18}$$

Proof. Since the sequence $(\varepsilon_i)_{i=0,1,\dots}$ is independent of $\sigma(\bigcup_{t\geq 0}\mathscr{F}_t)$ and the two strategies $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$ are measurable with respect to this σ-field, we get $\mathbb{E}[\,\varepsilon_k\xi_k\eta_k\,] = \frac{1}{2}\mathbb{E}[\,\xi_k\eta_k\,]$. Hence,

$$\begin{aligned}
\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,] - X_0 S_0^0 \quad &= \quad \mathbb{E}\Big[\,\sum_{k=0}^N \Big(\frac{1}{2}\xi_k^2 - S_{t_k}^{\boldsymbol{\xi},\boldsymbol{\eta}}\xi_k + \varepsilon_k\xi_k\eta_k + \theta\xi_k^2\Big)\,\Big] \\
&= \quad \mathbb{E}\Big[\,\sum_{k=0}^N \Big(\frac{1}{2}\xi_k^2 + \frac{1}{2}\xi_k\eta_k - \xi_k\Big(S_{t_k}^0 - \sum_{m=0}^{k-1}(\xi_m + \eta_m)G(t_k - t_m)\Big) + \theta\xi_k^2\Big)\,\Big] \\
&= \quad \mathbb{E}\Big[\, -\sum_{k=0}^N \xi_k S_{t_k}^0 + \frac{1}{2}\sum_{k=0}^N \xi_k^2 + \sum_{k=0}^N \xi_k \sum_{m=0}^{k-1}\xi_m G(t_k - t_m) \\
&\qquad\qquad + \sum_{k=0}^N \Big(\xi_k\Big(\frac{1}{2}\eta_k + \sum_{m=0}^{k-1}\eta_m G(t_k - t_m)\Big) + \theta\xi_k^2\Big)\,\Big].
\end{aligned}$$

Since each ξ_k is $\mathscr{F}_{t_k}$-measurable and S^0 is a martingale, we get from condition (b) in Definition 2.1 that

$$\mathbb{E}\Big[\,\sum_{k=0}^N \xi_k S_{t_k}^0\,\Big] = \mathbb{E}\Big[\,\sum_{k=0}^N \xi_k S_T^0\,\Big] = X_0 \mathbb{E}[\,S_T^0\,] = X_0 S_0^0.$$

Moreover,

$$\frac{1}{2}\sum_{k=0}^{N}\xi_k^2+\sum_{k=0}^{N}\xi_k\sum_{m=0}^{k-1}\xi_m G(t_k-t_m)=\frac{1}{2}\sum_{k,m=0}^{N}\xi_k\xi_m G(|t_k-t_m|)=\frac{1}{2}\boldsymbol{\xi}^\top\Gamma\boldsymbol{\xi},$$

and

$$\sum_{k=0}^{N}\xi_k\Big(\frac{1}{2}\eta_k+\sum_{m=0}^{k-1}\eta_m G(t_k-t_m)\Big)=\boldsymbol{\xi}^\top\widetilde{\Gamma}\boldsymbol{\eta}.$$

Putting everything together yields the assertion. $\square$

We will use the convention of saying that an $n\times n$-matrix A is *positive definite* when $\boldsymbol{x}^\top A\boldsymbol{x}>0$ for all nonzero $\boldsymbol{x}\in\mathbb{R}^n$, even when A is not necessarily symmetric. Clearly, for a positive definite matrix A there is no nonzero $\boldsymbol{x}\in\mathbb{R}^n$ for which $A\boldsymbol{x}=\mathbf{0}$, and so A is invertible. Moreover, writing a given nonzero $\boldsymbol{x}\in\mathbb{R}^n$ as $\boldsymbol{x}=A\boldsymbol{y}$ for $\boldsymbol{y}=A^{-1}\boldsymbol{x}\neq\mathbf{0}$, we see that $\boldsymbol{x}^\top A^{-1}\boldsymbol{x}=\boldsymbol{y}^\top A^\top\boldsymbol{y}=\boldsymbol{y}^\top A\boldsymbol{y}>0$. So the inverse of a positive definite matrix is also positive definite. Recall that we assume (2) throughout this paper.

Lemma 3.2. *The matrices Γ_θ, $\widetilde{\Gamma}$, $\Gamma_\theta+\widetilde{\Gamma}$, $\Gamma_\theta-\widetilde{\Gamma}$ are positive definite for all $\theta\geq 0$. In particular, all terms in* (7) *are well-defined and the denominators in* (7) *are strictly positive.*

Proof. That Γ is positive definite follows directly from (2). Therefore, for nonzero $\boldsymbol{x}\in\mathbb{R}^{N+1}$,

$$0<\boldsymbol{x}^\top\Gamma\boldsymbol{x}=\boldsymbol{x}^\top(\widetilde{\Gamma}+\widetilde{\Gamma}^\top)\boldsymbol{x}=\boldsymbol{x}^\top\widetilde{\Gamma}\boldsymbol{x}+\boldsymbol{x}^\top\widetilde{\Gamma}^\top\boldsymbol{x}=2\boldsymbol{x}^\top\widetilde{\Gamma}\boldsymbol{x},$$

which shows that the matrix $\widetilde{\Gamma}$ is positive definite. Next, $\Gamma-\widetilde{\Gamma}=\widetilde{\Gamma}^\top$ and so this matrix is also positive definite. Clearly, the sum of two positive definite matrices is also positive definite, which shows that $\Gamma_\theta+\widetilde{\Gamma}=\Gamma+\widetilde{\Gamma}+2\theta\,\mathrm{Id}$ and $\Gamma_\theta-\widetilde{\Gamma}=\Gamma-\widetilde{\Gamma}+2\theta\,\mathrm{Id}$ are positive definite for $\theta\geq 0$. $\square$

Lemma 3.3. *For given time grid $\mathbb{T}$ and initial values X_0 and Y_0, there exists at most one Nash equilibrium in the class $\mathscr{X}(X_0,\mathbb{T})\times\mathscr{X}(Y_0,\mathbb{T})$.*

Proof. We assume by way of contradiction that there exist two distinct Nash equilibria $(\boldsymbol{\xi}^0,\boldsymbol{\eta}^0)$ and $(\boldsymbol{\xi}^1,\boldsymbol{\eta}^1)$ in $\mathscr{X}(X_0,\mathbb{T})\times\mathscr{X}(Y_0,\mathbb{T})$. Here, the fact that the two Nash equilibria are distinct means that they are not $\mathbb{P}$-a.s. equal. Then we define for $\alpha\in[0,1]$

$$\boldsymbol{\xi}^\alpha:=\alpha\boldsymbol{\xi}^1+(1-\alpha)\boldsymbol{\xi}^0\qquad\text{and}\qquad\boldsymbol{\eta}^\alpha:=\alpha\boldsymbol{\eta}^1+(1-\alpha)\boldsymbol{\eta}^0.$$

We furthermore let

$$f(\alpha):=\mathbb{E}\Big[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}^\alpha|\boldsymbol{\eta}^0)+\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}^\alpha|\boldsymbol{\xi}^0)+\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}^{1-\alpha}|\boldsymbol{\eta}^1)+\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}^{1-\alpha}|\boldsymbol{\xi}^1)\,\Big].$$

Since according to Lemma 3.2 the matrix Γ_θ is positive definite, the functional

$$\boldsymbol{\xi}\longmapsto\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,]=\mathbb{E}\Big[\,\frac{1}{2}\boldsymbol{\xi}^\top\Gamma_\theta\boldsymbol{\xi}+\boldsymbol{\xi}^\top\widetilde{\Gamma}\boldsymbol{\eta}\,\Big]$$

is strictly convex with respect to $\boldsymbol{\xi}$. Since the two Nash equilibria $(\boldsymbol{\xi}^0,\boldsymbol{\eta}^0)$ and $(\boldsymbol{\xi}^1,\boldsymbol{\eta}^1)$ are distinct, $f(\alpha)$ must also be strictly convex in α and have its unique minimum in $\alpha=0$. That is,

$$f(\alpha)>f(0)\quad\text{for }\alpha>0.\tag{19}$$

It follows that

$$\lim_{h\downarrow 0}\frac{f(h)-f(0)}{h}=\frac{df(\alpha)}{d\alpha}\Big|_{\alpha=0+}\geq 0. \tag{20}$$

Next, by the symmetry of Γ_θ,

$$\begin{aligned}\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}^\alpha|\boldsymbol{\eta})\,] &= \mathbb{E}\bigg[\frac{1}{2}\alpha^2(\boldsymbol{\xi}^1)^\top\Gamma_\theta\boldsymbol{\xi}^1+\alpha(1-\alpha)(\boldsymbol{\xi}^1)^\top\Gamma_\theta\boldsymbol{\xi}^0+\frac{1}{2}(1-\alpha)^2(\boldsymbol{\xi}^0)^\top\Gamma_\theta\boldsymbol{\xi}^0\\ &\qquad +\alpha(\boldsymbol{\xi}^1)^\top\widetilde{\Gamma}\boldsymbol{\eta}+(1-\alpha)(\boldsymbol{\xi}^0)^\top\widetilde{\Gamma}\boldsymbol{\eta}\bigg].\end{aligned}$$

Therefore,

$$\frac{d}{d\alpha}\Big|_{\alpha=0+}\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}^\alpha|\boldsymbol{\eta})\,]=\mathbb{E}\Big[\,(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma_\theta\boldsymbol{\xi}^0+(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\widetilde{\Gamma}\boldsymbol{\eta}\Big].$$

Hence, it follows that

$$\begin{aligned}&\frac{d}{d\alpha}\Big|_{\alpha=0+}f(\alpha)\\ &= \mathbb{E}\bigg[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma_\theta\boldsymbol{\xi}^0+(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\widetilde{\Gamma}\boldsymbol{\eta}^0+(\boldsymbol{\xi}^0-\boldsymbol{\xi}^1)^\top\Gamma_\theta\boldsymbol{\xi}^1+(\boldsymbol{\xi}^0-\boldsymbol{\xi}^1)^\top\widetilde{\Gamma}\boldsymbol{\eta}^1\\ &\qquad +(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\Gamma_\theta\boldsymbol{\eta}^0+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\widetilde{\Gamma}\boldsymbol{\xi}^0+(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)^\top\Gamma_\theta\boldsymbol{\eta}^1+(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)^\top\widetilde{\Gamma}\boldsymbol{\xi}^1\bigg]\\ &= -\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma_\theta(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\Gamma_\theta(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big]\\ &\quad +\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\widetilde{\Gamma}(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)+(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\widetilde{\Gamma}^\top(\boldsymbol{\eta}^0-\boldsymbol{\eta}^1)\Big]\\ &= -\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma_\theta(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\Gamma_\theta(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big]-\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big].\end{aligned}$$

Now,

$$\begin{aligned}&(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)+\frac{1}{2}\Big((\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma_\theta(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\Gamma_\theta(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big)\\ &\qquad\geq\frac{1}{2}\Big((\boldsymbol{\xi}^1-\boldsymbol{\xi}^0+\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\Gamma(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0+\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big)\geq 0.\end{aligned}$$

Thus, and because the two Nash equilibria $(\boldsymbol{\xi}^0,\boldsymbol{\eta}^0)$ and $(\boldsymbol{\xi}^1,\boldsymbol{\eta}^1)$ are distinct, we have

$$\frac{d}{d\alpha}\Big|_{\alpha=0+}f(\alpha)\leq-\frac{1}{2}\mathbb{E}\Big[(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)^\top\Gamma(\boldsymbol{\xi}^1-\boldsymbol{\xi}^0)+(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)^\top\Gamma(\boldsymbol{\eta}^1-\boldsymbol{\eta}^0)\Big]<0,$$

which contradicts (20). Therefore, there can exist at most one Nash equilibrium in the class $\mathscr{X}(X_0,\mathbb{T})\times\mathscr{X}(Y_0,\mathbb{T})$. $\square$

Now let us introduce the class

$$\mathscr{X}_{\mathrm{det}}(Z_0,\mathbb{T}):=\Big\{\boldsymbol{\zeta}\in\mathscr{X}(Z_0,\mathbb{T})\,\Big|\,\boldsymbol{\zeta}\text{ is deterministic}\Big\}$$

of deterministic strategies in $\mathscr{X}(Z_0,\mathbb{T})$. A Nash equilibrium in the class $\mathscr{X}_{\mathrm{det}}(X_0,\mathbb{T})\times\mathscr{X}_{\mathrm{det}}(Y_0,\mathbb{T})$ is defined in the same way as in Definition 2.5.

Lemma 3.4. *A Nash equilibrium in the class $\mathscr{X}_{\mathrm{det}}(X_0,\mathbb{T})\times\mathscr{X}_{\mathrm{det}}(Y_0,\mathbb{T})$ of deterministic strategies is also a Nash equilibrium in the class $\mathscr{X}(X_0,\mathbb{T})\times\mathscr{X}(Y_0,\mathbb{T})$ of adapted strategies.*

Proof. Assume that $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)$ is a Nash equilibrium in the class $\mathscr{X}_{\mathrm{det}}(X_0,\mathbb{T})\times\mathscr{X}_{\mathrm{det}}(Y_0,\mathbb{T})$ of deterministic strategies. We need to show that $\boldsymbol{\xi}^*$ minimizes $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,]$ and $\boldsymbol{\eta}^*$ minimizes $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,]$ in the respective classes $\mathscr{X}(X_0,\mathbb{T})$ and $\mathscr{X}(Y_0,\mathbb{T})$ of adapted strategies. To this end, let $\boldsymbol{\xi}\in\mathscr{X}(X_0,\mathbb{T})$ be given. We define $\overline{\boldsymbol{\xi}}\in\mathscr{X}_{\mathrm{det}}(X_0,\mathbb{T})$ by $\overline{\xi}_k=\mathbb{E}[\,\xi_k\,]$ for $k=0,1,\dots,N$.

Applying Jensen's inequality to the convex function $\mathbb{R}^{N+1}\ni\boldsymbol{x}\mapsto\boldsymbol{x}^\top\Gamma_\theta\boldsymbol{x}$, we obtain

$$\begin{aligned}\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,] &= \mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top\Gamma_\theta\boldsymbol{\xi}+\boldsymbol{\xi}^\top\widetilde{\Gamma}\boldsymbol{\eta}^*\Big]=\mathbb{E}\Big[\frac{1}{2}\boldsymbol{\xi}^\top\Gamma_\theta\boldsymbol{\xi}\Big]+\overline{\boldsymbol{\xi}}^\top\widetilde{\Gamma}\boldsymbol{\eta}^*\\ &\geq \frac{1}{2}\overline{\boldsymbol{\xi}}^\top\Gamma_\theta\overline{\boldsymbol{\xi}}+\overline{\boldsymbol{\xi}}^\top\widetilde{\Gamma}\boldsymbol{\eta}^*=\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\overline{\boldsymbol{\xi}}|\boldsymbol{\eta}^*)\,]\\ &\geq \mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}^*|\boldsymbol{\eta}^*)\,].\end{aligned}$$

This shows that $\boldsymbol{\xi}^*$ minimizes $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta}^*)\,]$ over $\boldsymbol{\xi}\in\mathscr{X}(X_0,\mathbb{T})$. One can show analogously that $\boldsymbol{\eta}^*$ minimizes $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,]$ over $\boldsymbol{\eta}\in\mathscr{X}(Y_0,\mathbb{T})$, which completes the proof. □

Remark 3.5. Before proving Theorem 2.6, we briefly explain how to derive heuristically the explicit form (8) of the equilibrium strategies. By Lemma 3.1 and the method of Lagrange multipliers, a necessary condition for $(\boldsymbol{\xi}^*,\boldsymbol{\eta}^*)$ to be a Nash equilibrium in $\mathscr{X}_{\mathrm{det}}(X_0,\mathbb{T})\times\mathscr{X}_{\mathrm{det}}(Y_0,\mathbb{T})$ is the existence of $\alpha,\beta\in\mathbb{R}$, such that

$$\begin{cases}\Gamma_\theta\boldsymbol{\xi}^*+\widetilde{\Gamma}\boldsymbol{\eta}^*=\alpha\mathbf{1};\\ \Gamma_\theta\boldsymbol{\eta}^*+\widetilde{\Gamma}\boldsymbol{\xi}^*=\beta\mathbf{1}.\end{cases}\tag{21}$$

By adding the equations in (21) we obtain

$$(\Gamma_\theta+\widetilde{\Gamma})(\boldsymbol{\xi}^*+\boldsymbol{\eta}^*)=(\alpha+\beta)\mathbf{1}.\tag{22}$$

By Lemma 3.2, the matrix $\Gamma_\theta+\widetilde{\Gamma}$ is positive definite and hence invertible, so that (22) can be solved for $\boldsymbol{\xi}^*+\boldsymbol{\eta}^*$. Since we must also have $\mathbf{1}^\top(\boldsymbol{\xi}^*+\boldsymbol{\eta}^*)=X_0+Y_0$, we obtain

$$\boldsymbol{\xi}^*+\boldsymbol{\eta}^*=\frac{(X_0+Y_0)}{\mathbf{1}^\top(\Gamma_\theta+\widetilde{\Gamma})^{-1}\mathbf{1}}(\Gamma_\theta+\widetilde{\Gamma})^{-1}\mathbf{1}=(X_0+Y_0)\boldsymbol{v}.$$

Similarly, by subtracting the two equations in (21) yields

$$(\Gamma_\theta-\widetilde{\Gamma})(\boldsymbol{\xi}^*-\boldsymbol{\eta}^*)=(\alpha-\beta)\mathbf{1}.$$

It follows again from Lemma 3.2 that $(\Gamma_\theta-\widetilde{\Gamma})$ is invertible, and so we have

$$\boldsymbol{\xi}^*-\boldsymbol{\eta}^*=\frac{(X_0-Y_0)}{\mathbf{1}^T(\Gamma_\theta-\widetilde{\Gamma})^{-1}\mathbf{1}}(\Gamma_\theta-\widetilde{\Gamma})^{-1}\mathbf{1}=(X_0-Y_0)\boldsymbol{w}.$$

Thus, $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ ought to be given by (8).

Proof of Theorem 2.6. By Lemmas 3.3 and 3.4 all we need to show is that (8) defines a Nash equilibrium in the class $\mathscr{X}_{\det}(X_0,\mathbb{T})\times\mathscr{X}_{\det}(Y_0,\mathbb{T})$ of deterministic strategies. For $(\boldsymbol{\xi},\boldsymbol{\eta})\in\mathscr{X}_{\det}(X_0,\mathbb{T})\times\mathscr{X}_{\det}(Y_0,\mathbb{T})$ we have

$$\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,]=\frac{1}{2}\boldsymbol{\xi}^\top\Gamma_\theta\boldsymbol{\xi}+\boldsymbol{\xi}^\top\widetilde{\Gamma}\boldsymbol{\eta}. \tag{23}$$

Therefore minimizing $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\xi}|\boldsymbol{\eta})\,]$ over $\boldsymbol{\xi}\in\mathscr{X}_{\det}(X_0,\mathbb{T})$ is equivalent to the minimization of the quadratic form on the right-hand side of (23) over $\boldsymbol{\xi}\in\mathbb{R}^{N+1}$ under the constraint $\mathbf{1}^\top\boldsymbol{\xi}=X_0$.

Now we prove that the strategies $\boldsymbol{\xi}^*$ and $\boldsymbol{\eta}^*$ given by (8) are indeed optimal. We have

$$\Gamma_\theta\boldsymbol{\xi}^*+\widetilde{\Gamma}\boldsymbol{\eta}^*=\frac{1}{2}(X_0+Y_0)(\Gamma_\theta+\widetilde{\Gamma})\boldsymbol{v}+\frac{1}{2}(X_0-Y_0)(\Gamma_\theta-\widetilde{\Gamma})\boldsymbol{w}=\mu\mathbf{1}, \tag{24}$$

where

$$\mu=\frac{(X_0+Y_0)}{2\mathbf{1}^\top(\Gamma_\theta+\widetilde{\Gamma})\mathbf{1}}+\frac{(X_0-Y_0)}{2\mathbf{1}^\top(\Gamma_\theta-\widetilde{\Gamma})\mathbf{1}}.$$

Now let $\boldsymbol{\xi}\in\mathscr{X}_{\det}(X_0,\mathbb{T})$ be arbitrary and define $\boldsymbol{\zeta}:=\boldsymbol{\xi}-\boldsymbol{\xi}^*$. Then we have $\boldsymbol{\zeta}^\top\mathbf{1}=0$. Hence, by the symmetry of Γ_θ,

$$\begin{aligned}\frac{1}{2}\boldsymbol{\xi}^\top\Gamma_\theta\boldsymbol{\xi}+\boldsymbol{\xi}^\top\widetilde{\Gamma}\boldsymbol{\eta}^* &= \frac{1}{2}(\boldsymbol{\xi}^*)^\top\Gamma_\theta\boldsymbol{\xi}^*+\frac{1}{2}\boldsymbol{\zeta}^\top\Gamma_\theta\boldsymbol{\zeta}+\boldsymbol{\zeta}^\top\Gamma_\theta\boldsymbol{\xi}^*+(\boldsymbol{\xi}^*)^\top\widetilde{\Gamma}\boldsymbol{\eta}^*+\boldsymbol{\zeta}^\top\widetilde{\Gamma}\boldsymbol{\eta}^*\\ &= \frac{1}{2}(\boldsymbol{\xi}^*)^\top\Gamma_\theta\boldsymbol{\xi}^*+(\boldsymbol{\xi}^*)^\top\widetilde{\Gamma}\boldsymbol{\eta}^*+\frac{1}{2}\boldsymbol{\zeta}^\top\Gamma_\theta\boldsymbol{\zeta}+\mu\boldsymbol{\zeta}^\top\mathbf{1}\\ &\geq \frac{1}{2}(\boldsymbol{\xi}^*)^\top\Gamma_\theta\boldsymbol{\xi}^*+(\boldsymbol{\xi}^*)^\top\widetilde{\Gamma}\boldsymbol{\eta}^*,\end{aligned}$$

where in the last step we have used that Γ_θ is positive definite and that $\boldsymbol{\zeta}^\top\mathbf{1}=0$. Therefore $\boldsymbol{\xi}^*$ minimizes (23) in the class $\mathscr{X}_{\det}(X_0,\mathbb{T})$ for $\boldsymbol{\eta}=\boldsymbol{\eta}^*$. In the same way, one shows that $\boldsymbol{\eta}^*$ minimizes $\mathbb{E}[\,\mathscr{C}_{\mathbb{T}}(\boldsymbol{\eta}|\boldsymbol{\xi}^*)\,]$ over $\boldsymbol{\eta}\in\mathscr{X}_{\det}(X_0,\mathbb{T})$. □

3.2 Proof of Propositions 2.7, 2.11, and 2.12

Recall that in this section $G(t)=\gamma+\lambda e^{-\rho t}$ for constants $\lambda,\rho>0$ and $\gamma\geq 0$.

Proof of Proposition 2.7. We need to compute the inverse of the matrix $\Gamma_\theta-\widetilde{\Gamma}$. Setting $\kappa:=2\theta/\lambda+\frac{1}{2}$ and $a:=e^{-\rho T}$, we have

$$\Gamma_\theta-\widetilde{\Gamma}=\lambda\begin{pmatrix}\kappa & a^{\frac{1}{N}} & a^{\frac{2}{N}} & \cdots & a^{\frac{N-1}{N}} & a\\ 0 & \kappa & a^{\frac{1}{N}} & \cdots & a^{\frac{N-2}{N}} & a^{\frac{N-1}{N}}\\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \kappa & a^{\frac{1}{N}}\\ 0 & \cdots & \cdots & \cdots & 0 & \kappa\end{pmatrix}.$$

It is easy to verify that the inverse of this matrix is given by

$$\Pi_N := \frac{1}{\lambda}\begin{pmatrix} \frac{1}{\kappa} & \frac{-a^{\frac{1}{N}}}{\kappa^2} & \frac{-a^{\frac{2}{N}}(\kappa-1)}{\kappa^3} & \cdots & \frac{-a^{\frac{N-1}{N}}(\kappa-1)^{N-2}}{\kappa^N} & \frac{-a^{\frac{N}{N}}(\kappa-1)^{N-1}}{\kappa^{N+1}} \\ 0 & \frac{1}{\kappa} & \frac{-a^{\frac{1}{N}}}{\kappa^2} & \cdots & \frac{-a^{\frac{N-2}{N}}(\kappa-1)^{N-3}}{\kappa^{N-1}} & \frac{-a^{\frac{N-1}{N}}(\kappa-1)^{N-2}}{\kappa^N} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{1}{\kappa} & \frac{-a^{\frac{1}{N}}}{\kappa^2} \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{1}{\kappa} \end{pmatrix}. \tag{25}$$

Let us denote by $\boldsymbol{u} = (u_1, u_2, \dots, u_{N+1}) \in \mathbb{R}^{N+1}$ the vector $\lambda\Pi_N\mathbf{1}$. Then we have $u_{N+1} = \frac{1}{\kappa}$ and, for $n = 1, \dots, N$, $u_n = u_{n+1} - a^{(N+1-n)/N}(\kappa-1)^{N-n}/\kappa^{N+2-n}$. That is,

$$\begin{aligned} u_n &= \frac{1}{\kappa} - \frac{a^{\frac{1}{N}}}{\kappa^2}\sum_{m=n}^{N}\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N-m} = \frac{1}{\kappa} - \frac{a^{\frac{1}{N}}}{\kappa^2}\sum_{k=0}^{N-n}\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{k} \\ &= \frac{1}{\kappa}\Big[1 - \frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}}) + a^{\frac{1}{N}}} + (-1)^{N+1-n}\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}}) + a^{\frac{1}{N}}}\Big(\frac{a^{\frac{1}{N}}(1-\kappa)}{\kappa}\Big)^{N+1-n}\Big]. \end{aligned} \tag{26}$$

When $\theta = 0$, we have

$$u_n = 2\Big[1 - \frac{2a^{\frac{1}{N}}}{1+a^{\frac{1}{N}}} + (-1)^{N+1-n}\frac{2a^{\frac{N+2-n}{N}}}{1+a^{\frac{1}{N}}}\Big]. \tag{27}$$

Since $a < 1$, we have

$$0 \le 1 - \frac{2a^{\frac{1}{N}}}{1+a^{\frac{1}{N}}} < 1 - a^{\frac{1}{N}} \longrightarrow 0 \qquad \text{as } N \uparrow \infty.$$

On the other hand, we have

$$\frac{2a^{\frac{N+2-n}{N}}}{1+a^{\frac{1}{N}}} \ge a^{\frac{N+2-n}{N}} \ge a^{\frac{N+1}{N}} \longrightarrow a \qquad \text{as } N \uparrow \infty.$$

Therefore, the signs of u_n will alternate as soon as N is large enough to have $1 - a^{\frac{1}{N}} < a^{\frac{N+1}{N}}$. This proves part (a). As for part (b), since the expression (26) is continuous in κ, the signs of u_n will still alternate if, for fixed $N \ge N_0$, we take κ slightly larger than $1/2$. (Note however that the term $(1-\kappa)^N/\kappa^N$ tends to zero faster than $1 - a^{\frac{1}{N}}$, so we cannot get this result uniformly in N). $\square$

Lemma 3.6. *Let Π_N be as in* (25) *and let us denote by $\boldsymbol{u}^{(N)} = (u_1^{(N)}, u_2^{(N)}, \dots, u_{N+1}^{(N)}) \in \mathbb{R}^{N+1}$ the vector $\lambda\Pi_N\mathbf{1}$. When $n \in \{1, \dots, N+1\}$, then*

$$\sum_{m=1}^{n} u_m^{(N)} = \frac{1}{\kappa}\Bigg[n\Big(1 - \frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}}) + a^{\frac{1}{N}}}\Big) + \frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}}) + a^{\frac{1}{N}}}\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N+1-n}\frac{\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^n - 1}{\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa} - 1}\Bigg].$$

Proof. The assertion follows from (26) by noting that

$$\sum_{m=1}^{n}\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N+1-m} = \Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N+1-n}\frac{\big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\big)^{n}-1}{\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}-1}.$$

□

Proof of Proposition 2.11. Let Π_N and $\boldsymbol{u}^{(N)}$ be as in Lemma 3.6. We need to normalize the vector $\boldsymbol{u}^{(N)}$ with $\mathbf{1}^\top\lambda\Pi_N\mathbf{1} = \mathbf{1}^\top\boldsymbol{u}^{(N)}$ to get $\boldsymbol{w}^{(N)}$. Taking $n = N+1$ in Lemma 3.6 yields

$$\begin{aligned}\mathbf{1}^\top\lambda\Pi_N\mathbf{1} &= \sum_{n=1}^{N+1} u_n^{(N)}\\ &= \frac{1}{\kappa}\bigg[(N+1)\bigg(1-\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}}\bigg)+\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}}\cdot\frac{\big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\big)^{N+1}-1}{\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}-1}\bigg].\end{aligned}$$

Note that

$$(N+1)\bigg(1-\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}}\bigg) \longrightarrow -\kappa\log a = \kappa\rho T \qquad \text{as } N\uparrow\infty. \tag{28}$$

Moreover, $\kappa \geq 1/2$ implies that $|\kappa-1|/\kappa \leq 1$ with equality if and only if $\kappa = 1/2$. We therefore get that for $\kappa = 1/2$, which is the same as $\theta = 0$,

$$\begin{aligned}\lim_{N\uparrow\infty}\mathbf{1}^\top\lambda\Pi_{2N}\mathbf{1} &= \rho T + \frac{1+a}{2\kappa} = \rho T + a + 1,\\ \lim_{N\uparrow\infty}\mathbf{1}^\top\lambda\Pi_{2N+1}\mathbf{1} &= \rho T + \frac{1-a}{2\kappa} = \rho T - a + 1.\end{aligned}\tag{29}$$

For $\kappa > 1/2$, we have

$$\lim_{N\uparrow\infty}\mathbf{1}^\top\lambda\Pi_N\mathbf{1} = \rho T + 1. \tag{30}$$

The assertions now follow easily by taking limits in (27) and (26). □

Proof of Proposition 2.12. Let $n_t := \lceil Nt/T\rceil$. Then, with the notation introduced in the proof of Proposition 2.11,

$$W_t^{(N)} = 1 - \frac{1}{\mathbf{1}^\top\lambda\Pi_N\mathbf{1}}\sum_{k=1}^{n_t} u_k^{(N)}.$$

For $\theta > 0$ and $t < T$ it follows from Lemma 3.6 that

$$\Big(\frac{a^{\frac{1}{N}}(\kappa-1)}{\kappa}\Big)^{N+1-n_t} \longrightarrow 0 \qquad \text{as } N\uparrow\infty.$$

Therefore, with (28),

$$\lim_{N\uparrow\infty}\sum_{k=1}^{n_t} u_k^{(N)} = \frac{1}{\kappa}\lim_{N\uparrow\infty} n_t\bigg(1-\frac{a^{\frac{1}{N}}}{\kappa(1-a^{\frac{1}{N}})+a^{\frac{1}{N}}}\bigg) = \rho t.$$

The assertion now follows with (30). □

3.3 Proof of Theorem 2.8

Ours proof relies on results for so-called M-matrices stated in the book [5] by Berman and Plemmons. We first introduce some notations. When A is a matrix or vector, we will write

(a) $A \geq 0$ if each entry of A is nonnegative;

(b) $A > 0$ if $A \geq 0$ and at least one entry is strictly positive;

(c) $A \gg 0$ if each entry of A is strictly positive.

Definition 3.7 (Definition 1.2 in Chapter 6 of [5])**.** A matrix $A \in \mathbb{R}^{n\times n}$ is called a nonsingular *M-matrix* if it is of the form $A = s\,\mathrm{Id} - B$, where the matrix $B \in \mathbb{R}^{n\times n}$ satisfies $B \geq 0$ and the parameter $s > 0$ is strictly larger than the spectral radius of B.

Also recall that a matrix $A \in \mathbb{R}^{n\times n}$ is called a *Z-matrix* if all its off-diagonal elements are nonpositive. Berman and Plemmons [5] give 50 equivalent characterizations of the fact that a given Z-matrix is a nonsingular M-matrix. We will need three of them here and summarize them in the following statement.

Theorem 3.8 (From Theorem 2.3 in Chapter 6 of [5])**.** *For a Z-matrix $A \in \mathbb{R}^{n\times n}$, the following conditions are equivalent.*

(a) *A is a nonsingular M-matrix;*

(b) *All the leading principal minors of A are positive.*

(c) *A is inverse-positive; that is, A^{-1} exists and $A^{-1} \geq 0$.*

(d) $A + \alpha\,\mathrm{Id}$ *is nonsingular for all $\alpha \geq 0$.*

We start with the following auxiliary lemma.

Lemma 3.9. *A triangular Z-matrix $A \in \mathbb{R}^{n\times n}$ with positive diagonal is an M-matrix.*

Proof. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & a_{nn} \end{pmatrix}$$

be an upper triangular Z-matrix with positive diagonal. Then all of its leading principle minors are positive:

$$A_{[k]} = \prod_{i=1}^{k} a_{ii} > 0, \text{ for } k \in 1, 2, \ldots, N.$$

By Theorem 3.8 (b), A is an M-matrix. □

It will be convenient to define the matrices

$$\Phi_{ij} := e^{-\rho|t_{i-1}-t_{j-1}|} \qquad \text{and} \qquad \Psi_{ij} := 1 \tag{31}$$

for $i, j = 1, \ldots, N+1$. Recalling that $G(t) = \lambda e^{-\rho t} + \gamma$, we then have $\Gamma = \lambda\Phi + \gamma\Psi$ and $\Gamma_\theta = \lambda\Phi + \gamma\Psi + 2\theta\,\mathrm{Id}$. Moreover, for any matrix A we let

$$\widetilde{A}_{ij} := \begin{cases} A_{ij} & \text{if } i > j, \\ \frac{1}{2}A_{ij} & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Note that this notation is consistent with (6), and we get $\widetilde{\Gamma} = \lambda\widetilde{\Phi} + \gamma\widetilde{\Phi}$. We finally define

$$\widehat{\Phi} := \widetilde{\Phi} + \frac{1}{2}\,\mathrm{Id}.$$

Lemma 3.10. *For $\alpha \geq 0$, the inverse of the matrix $\widehat{\Phi} + \alpha\Phi$ is given by*

$$\begin{pmatrix} \beta & -a^{\frac{1}{N}}\mu\beta^2 & -a^{\frac{2}{N}}\mu\beta^3 & \cdots & -a^{\frac{N-1}{N}}\mu\beta^N & -\frac{a\alpha}{1+\alpha}\beta^N \\ -a^{\frac{1}{N}}\beta & (1+(1-a^{\frac{4}{N}})\alpha)\beta^2 & -a^{\frac{1}{N}}\mu\nu\beta^3 & \cdots & -a^{\frac{N-2}{N}}\mu\nu\beta^N & -a^{\frac{N-1}{N}}\mu\beta^N \\ 0 & -a^{\frac{1}{N}}\beta & (1+(1-a^{\frac{4}{N}})\alpha)\beta^2 & \cdots & -a^{\frac{N-3}{N}}\mu\nu\beta^{N-1} & -a^{\frac{N-2}{N}}\mu\beta^{N-1} \\ 0 & 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & -a^{\frac{1}{N}}\beta & (1+(1-a^{\frac{4}{N}})\alpha)\beta^2 & -a^{\frac{1}{N}}\mu\beta^2 \\ 0 & \cdots & \cdots & 0 & -a^{\frac{1}{N}}\beta & \beta \end{pmatrix},$$

where

$$\beta = \big(1+(1-a^{\frac{2}{N}})\alpha\big)^{-1}, \qquad \mu = (1-a^{\frac{2}{N}})\alpha, \qquad \nu = (1-a^{\frac{2}{N}})(1+\alpha).$$

Proof. Let the matrix in the statement be denoted by P. We rewrite P as

$$P_{ij} = \begin{cases} \beta, & \text{if } i = j = 1 \text{ or } i = j = N+1; \\ (1+(1-a^{\frac{4}{N}})\alpha)\beta^2, & \text{if } i = j \in \{2, \ldots, N\}; \\ -a^{\frac{1}{N}}\beta, & \text{if } i - j = 1; \\ -a^{\frac{j-i}{N}}\beta^{j-i+2}\mu\nu, & \text{if } j - i \in \{1, \ldots, N-2\} \text{ and } i \neq 1 \text{ and } j \neq N+1; \\ -a^{\frac{j-i}{N}}\beta^{j-i+1}\mu, & \text{if } j - i \in \{1, \ldots, N-1\} \text{ and either } i = 1 \text{ or } j = N+1; \\ -\frac{a\alpha}{1+\alpha}\beta^N, & \text{if } i = 1 \text{ and } j = N+1, \\ 0, & \text{if } i \geq j+2. \end{cases}$$

On the other hand, the matrix $\widehat{\Phi} + \alpha\Phi$ can be written as

$$(\widehat{\Phi} + \alpha\Phi)_{ij} = \begin{cases} 1+\alpha, & \text{if } i = j; \\ \alpha a^{\frac{j-i}{N}}, & \text{if } i < j; \\ (1+\alpha)a^{\frac{i-j}{N}}, & \text{if } i > j. \end{cases}$$

Checking

$$\sum_{k=1}^{N+1} P_{ik}(\widehat{\Phi} + \alpha\Phi)_{kj} = \sum_{k=1}^{N+1} (\widehat{\Phi} + \alpha\Phi)_{ik}P_{kj} = \delta_{ij}$$

for all i and j completes the proof. $\square$

Let

$$\widehat{\Psi} := \widetilde{\Psi}^\top - \frac{1}{2}\,\mathrm{Id}.$$

Lemma 3.11. *The matrix $\Phi^{-1}\big(\widehat{\Phi} - \frac{\gamma}{\lambda}\widehat{\Psi}\big)$ is a Z-matrix and a nonsingular M-matrix.*

Proof. It was shown in [1, Theorem 3.4] that

$$\Phi^{-1} = \frac{1}{1-a^{\frac{2}{N}}}\begin{pmatrix} 1 & -a^{\frac{1}{N}} & 0 & \cdots & \cdots & 0 \\ -a^{\frac{1}{N}} & 1+a^{\frac{2}{N}} & -a^{\frac{1}{N}} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -a^{\frac{1}{N}} & 1+a^{\frac{2}{N}} & -a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & 0 & -a^{\frac{1}{N}} & 1 \end{pmatrix}. \tag{32}$$

The matrix $\widehat{\Phi} - \frac{\gamma}{\lambda}\widehat{\Psi}$ is equal to

$$\begin{pmatrix} 1 & -\frac{\gamma}{\lambda} & -\frac{\gamma}{\lambda} & \cdots & \cdots & -\frac{\gamma}{\lambda} \\ a^{\frac{1}{N}} & 1 & -\frac{\gamma}{\lambda} & \cdots & \cdots & -\frac{\gamma}{\lambda} \\ a^{\frac{2}{N}} & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a^{\frac{1}{N}} & 1 & -\frac{\gamma}{\lambda} \\ a & \cdots & \cdots & a^{\frac{2}{N}} & a^{\frac{1}{N}} & 1 \end{pmatrix}.$$

A straightforward computation now yields that the matrix $(1-a^{\frac{2}{N}})\Phi^{-1}\big(\widehat{\Phi} - \frac{\gamma}{\lambda}\widehat{\Psi}\big)$ is equal to

$$\begin{pmatrix} 1-a^{\frac{2}{N}} & -a^{\frac{1}{N}} - \frac{\gamma}{\lambda} & -(1-a^{\frac{1}{N}})\frac{\gamma}{\lambda} & -(1-a^{\frac{1}{N}})\frac{\gamma}{\lambda} & \cdots & -(1-a^{\frac{1}{N}})\frac{\gamma}{\lambda} \\ 0 & 1+a^{\frac{1}{N}}\frac{\gamma}{\lambda} & -a^{\frac{1}{N}} - (1-a^{\frac{1}{N}}+a^{\frac{2}{N}})\frac{\gamma}{\lambda} & -(1-a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} & \cdots & -(1-a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} \\ 0 & 0 & 1+a^{\frac{1}{N}}\frac{\gamma}{\lambda} & -a^{\frac{1}{N}} - (1-a^{\frac{1}{N}}+a^{\frac{2}{N}})\frac{\gamma}{\lambda} & \cdots & -(1-a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & -(1-a^{\frac{1}{N}})^2\frac{\gamma}{\lambda} \\ \vdots & \ddots & \ddots & 0 & 1+a^{\frac{1}{N}}\frac{\gamma}{\lambda} & -a^{\frac{1}{N}} - (1-a^{\frac{1}{N}}+a^{\frac{2}{N}})\frac{\gamma}{\lambda} \\ 0 & \cdots & \cdots & 0 & 0 & 1+a^{\frac{1}{N}}\frac{\gamma}{\lambda} \end{pmatrix},$$

which is an upper triangular Z-matrix with positive diagonal. By Lemma 3.9, $\Phi^{-1}\big(\widehat{\Phi} - \frac{\gamma}{\lambda}\widehat{\Psi}\big)$ is hence a nonsingular M-matrix.

□

Lemma 3.12. *For $\delta \geq 0$ the matrix $\Lambda_\delta := \Phi^{-1}\big(\widehat{\Phi} - \frac{\gamma}{\lambda}\widehat{\Psi}\big) + \delta\Phi^{-1}$ is a nonsingular M-matrix.*

Proof. For $\delta = 0$ the result follows from Lemma 3.11. So let us assume henceforth that $\delta > 0$. Note first that Λ_δ is a Z-matrix since both $\Phi^{-1}\big(\widehat{\Phi} - \frac{\gamma}{\lambda}\widehat{\Psi}\big)$ and Φ^{-1} are Z matrices by Lemma 3.11 and (32), respectively. Hence condition (d) of Theorem 3.8 will imply that Λ_δ is a nonsingular M-matrix as soon as we can show that $\Lambda_\delta + \alpha\,\mathrm{Id}$ is invertible for all $\alpha \geq 0$.

In a first step, we note that taking $\gamma = 0$ in Lemma 3.11 yields that $\Phi^{-1}\widehat{\Phi}$ is a nonsingular M-matrix. Hence $(\alpha\,\mathrm{Id} + \Phi^{-1}\widehat{\Phi})^{-1} \geq 0$ for all $\alpha \geq 0$. It follows that

$$\big(\widehat{\Phi} + \alpha\Phi\big)^{-1}\mathbf{1} = \big(\mathrm{Id} + (\alpha\Phi)^{-1}\widehat{\Phi}\big)^{-1}(\alpha\Phi)^{-1}\mathbf{1} = \Big(\alpha\,\mathrm{Id} + \Phi^{-1}\widehat{\Phi}\Big)^{-1}\Phi^{-1}\mathbf{1} > 0,$$

because by [1, Example 3.5],

$$\Phi^{-1}\mathbf{1} = \frac{1}{1 + a^{\frac{1}{N}}}\Big(1, 1 - a^{\frac{1}{N}}, \dots, 1 - a^{\frac{1}{N}}, 1\Big)^T \gg 0. \tag{33}$$

Since moreover $(\widehat{\Phi}+\alpha\Phi)^{-1}$ is a Z-matrix by Lemma 3.10, it follows that $(\widehat{\Phi}+\alpha\Phi)^{-1}$ is a diagonally dominant Z-matrix for all $\alpha \geq 0$.

In the next step, we show that the matrix

$$Q := (\widehat{\Phi} + \alpha\Phi)^{-1}\Big(\delta\,\mathrm{Id} - \frac{\gamma}{\lambda}\widehat{\Psi}\Big)$$

is a Z-matrix. Denoting again $P := (\widehat{\Phi} + \alpha\Phi)^{-1}$, we get

$$Q_{ij} = \delta P_{ij} - \frac{\gamma}{\lambda}\sum_{k=1}^{j-1} P_{ik},$$

with the convention that $\sum_{k=1}^{0} a_k = 0$. It follows that $Q_{ii} \geq 0$ for all i, because $P_{ii} \geq 0$ and $\frac{\gamma}{\lambda}\sum_{k=1}^{i-1} P_{ik} \leq 0$ by the fact that P is a Z-matrix. Since P is diagonally dominant, we have $\sum_{k=1}^{j-1} P_{ik} \geq 0$ for any $j > i$ and hence $Q_{ij} = \delta P_{ij} - \frac{\gamma}{\lambda}\sum_{k=1}^{j-1} P_{ik} \leq 0$ for $j > i$. Using the fact that $P_{ik} = 0$ for $k \leq i - 1$, we get that for $j < i$

$$Q_{ij} = \delta P_{ij} - \frac{\gamma}{\lambda}\sum_{k=1}^{j-1} P_{ik} = \delta P_{ij} \leq 0.$$

This shows that Q is a Z-matrix.

We show next that Q is a nonsingular M-matrix. To this end, we note first that the triangular matrix $\Big(\delta\,\mathrm{Id} - \frac{\gamma}{\lambda}\widehat{\Psi}\Big)$ is invertible under our assumption $\delta > 0$. As a matter of fact, an easy calculation verifies that its inverse is given by

$$\frac{1}{\delta}\begin{pmatrix} 1 & \sigma & \sigma(1+\sigma) & \cdots & \sigma(1+\sigma)^{N-2} & \sigma(1+\sigma)^{N-1} \\ 0 & 1 & \sigma & \sigma(1+\sigma) & \cdots & \sigma(1+\sigma)^{N-2} \\ 0 & 0 & 1 & \sigma & \cdots & \sigma(1+\sigma)^{N-3} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & \sigma \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \geq 0,$$

where $\sigma := \frac{\gamma}{\lambda\delta} > 0$. Hence,

$$Q^{-1} = \Big(\delta\,\mathrm{Id} - \frac{\gamma}{\lambda}\widehat{\Psi}\Big)^{-1}(\widehat{\Phi} + \alpha\Phi) \geq 0.$$

So Theorem 3.8 (c) shows that Q is a nonsingular M-matrix.

For the final step, we note first that Theorem 3.8 (d) implies that $\mathrm{Id} + Q$ is a nonsingular M-matrix. In particular, $(\mathrm{Id} + Q)^{-1}$ exists, and so we can define the matrix

$$\begin{aligned}(\mathrm{Id}+Q)^{-1}(\widehat{\Phi}+\alpha\Phi)^{-1}\Phi &= \Big(\delta\,\mathrm{Id}+\widehat{\Phi}+\alpha\Phi-\frac{\gamma}{\lambda}\widehat{\Psi}\Big)^{-1}\Phi\\ &= \Big(\mathrm{Id}+(\delta\Phi^{-1})^{-1}\Big(\Phi^{-1}\Big(\widehat{\Phi}-\frac{\gamma}{\lambda}\widehat{\Psi}\Big)+\alpha\,\mathrm{Id}\Big)\Big)^{-1}(\delta\Phi^{-1})^{-1}\\ &= \Big(\delta\Phi^{-1}+\Phi^{-1}\Big(\widehat{\Phi}-\frac{\gamma}{\lambda}\widehat{\Psi}\Big)+\alpha\,\mathrm{Id}\Big)^{-1}\\ &= (\Lambda_\delta+\alpha\,\mathrm{Id})^{-1}.\end{aligned}$$

This proves that $\Lambda_\delta + \alpha\,\mathrm{Id}$ is invertible and the proof is complete. □

Lemma 3.13. *Let A be an invertible matrix and suppose that $\alpha \in \mathbb{R}$ is such that $A + \alpha\Psi$ is invertible. Then the vector $A^{-1}\mathbf{1}$ is proportional to $(A + \alpha\Psi)^{-1}\mathbf{1}$.*

Proof. Note that $\Psi\boldsymbol{x}$ is proportional to $\mathbf{1}$ for any vector $\boldsymbol{x}$. Hence,

$$(A+\alpha\Psi)A^{-1}\mathbf{1} = (\mathrm{Id}+\alpha\Psi A^{-1})\mathbf{1} = (1+\beta)\mathbf{1}$$

for some constant β. Applying $(A + \alpha\Psi)^{-1}$ to both sides of this equation yields the result. □

We are now ready to prove Theorem 2.8. We will first prove that (c)⇔(a). Then we will show (c)⇔(b).

Proof of (c)⇒(a) *in Theorem 2.8.* We need to show that $\boldsymbol{v}$ has only nonnegative components for $\theta \geq \frac{\lambda+\gamma}{4}$. The vector $\boldsymbol{v}$ is proportional to $(\Gamma_\theta + \widetilde{\Gamma})^{-1}\mathbf{1}$. When setting

$$\delta := \frac{4\theta-(\lambda+\gamma)}{2\lambda} \geq 0,$$

we find that

$$\begin{aligned}\Gamma_\theta+\widetilde{\Gamma}-\gamma\Psi &= \lambda\Phi+\gamma\Psi+2\theta\,\mathrm{Id}+\lambda\widetilde{\Phi}+\gamma\widetilde{\Psi}-\gamma\Psi\\ &= \lambda\Phi+2\theta\,\mathrm{Id}+\lambda\widetilde{\Phi}-\gamma\widetilde{\Psi}^\top = \lambda\Phi+\lambda\Big(\widehat{\Phi}-\frac{\gamma}{\lambda}\widehat{\Psi}+\delta\,\mathrm{Id}\Big) = \lambda\Phi\big(\Lambda_\delta+\mathrm{Id}\big),\end{aligned}$$

and we know from Lemma 3.12 that the latter matrix is invertible. It therefore follows from Lemma 3.13 that $\boldsymbol{v}$ is proportional to

$$\big(\Phi(\Lambda_\delta+\mathrm{Id})\big)^{-1}\mathbf{1} = (\Lambda_\delta+\mathrm{Id})^{-1}\Phi^{-1}\mathbf{1}.$$

As noted in (33), we have $\Phi^{-1}\mathbf{1} \gg 0$. Moreover, Λ_δ, and hence $\Lambda_\delta + \mathrm{Id}$, are nonsingular M-matrices by Lemma 3.12 and Theorem 3.8 (d). Via Theorem 3.8 (c), these facts imply that $\big(\Phi(\Lambda_\delta + \mathrm{Id})\big)^{-1}\mathbf{1} \geq 0$ and in turn that $\boldsymbol{v} \geq 0$. □

Proof of (a)⇒(c) *in Theorem 2.8.* We consider the case $N = 1$. By definition, $\boldsymbol{v}$ is proportional to the vector

$$2\det(\Gamma_\theta + \widetilde{\Gamma})(\Gamma_\theta + \widetilde{\Gamma})^{-1}\mathbf{1} = \begin{pmatrix} \lambda(3-2a) + \gamma + 4\theta \\ \lambda(3-4a) - \gamma + 4\theta \end{pmatrix}.$$

Clearly, the first component of this vector is positive for all $a \in (0,1)$ and $\theta \geq 0$. By sending $a \uparrow 1$ one sees, however, that the second component is negative for $\theta < \theta^*$ and a sufficiently close to 1. Thus, we cannot have $\boldsymbol{v} \geq 0$ in this case. □

Proof of (c)⇒(b) *in Theorem 2.8.* We assume that $\theta \geq \theta^*$. Note that

$$\Big(\widetilde{\Psi}^\top + \frac{1}{2}\,\mathrm{Id}\Big)^{-1} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & 1 & \cdots & 1 & 1 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Letting $\kappa := \frac{1}{2} + \frac{2\theta}{\lambda} - \frac{\gamma}{2\lambda}$, it follows that

$$\begin{aligned}
&\frac{\gamma}{\lambda}\Big(\gamma\widetilde{\Psi}^\top + \frac{\gamma}{2}\,\mathrm{Id}\Big)^{-1}\Big(\lambda\widetilde{\Phi}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\,\mathrm{Id}\Big) \\
&= \Big(\widetilde{\Psi}^\top + \frac{1}{2}\,\mathrm{Id}\Big)^{-1}\Big(\widetilde{\Phi}^\top + \Big(\kappa - \frac{1}{2}\Big)\,\mathrm{Id}\Big) \\
&= \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 & -1 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}
\begin{pmatrix} \kappa & a^{\frac{1}{N}} & a^{\frac{2}{N}} & \cdots & a^{\frac{N-1}{N}} & a \\ 0 & \kappa & a^{\frac{1}{N}} & \cdots & a^{\frac{N-2}{N}} & a^{\frac{N-1}{N}} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \kappa & a^{\frac{1}{N}} \\ 0 & \cdots & \cdots & \cdots & 0 & \kappa \end{pmatrix}.
\end{aligned}$$

A straightforward computation yields that the preceding matrix product is equal to

$$\begin{pmatrix} \kappa & a^{\frac{1}{N}} - \kappa & (a^{\frac{1}{N}} - 1)a^{\frac{1}{N}} & \cdots & (a^{\frac{1}{N}} - 1)a^{\frac{N-2}{N}} & (a^{\frac{1}{N}} - 1)a^{\frac{N-1}{N}} \\ 0 & \kappa & a^{\frac{1}{N}} - \kappa & (a^{\frac{1}{N}} - 1)a^{\frac{1}{N}} & \cdots & (a^{\frac{1}{N}} - 1)a^{\frac{N-2}{N}} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & (a^{\frac{1}{N}} - 1)a^{\frac{1}{N}} \\ \vdots & \ddots & \ddots & \ddots & \kappa & a^{\frac{1}{N}} - \kappa \\ 0 & \cdots & \cdots & \cdots & 0 & \kappa \end{pmatrix}.$$

We have $\theta \geq \theta^*$ if and only if $\kappa \geq 1$. Therefore and since $a^{\frac{1}{N}} < 1$ for all $N \in \mathbb{N}$, the preceding is a Z-matrix. By Lemma 3.9, it is an M-matrix. Therefore, by Theorem 3.8, also the matrix

$$A := \mathrm{Id} + \Big(\gamma\widetilde{\Psi}^\top + \frac{\gamma}{2}\,\mathrm{Id}\Big)^{-1}\Big(\lambda\widetilde{\Phi}^\top + \Big(2\theta - \frac{\gamma}{2}\Big)\,\mathrm{Id}\Big)$$

is an M-matrix.

We have

$$\Big(\gamma\widetilde{\Psi}^\top + \frac{\gamma}{2}\,\mathrm{Id}\Big)A = \gamma\widetilde{\Psi}^\top + \lambda\widetilde{\Phi}^\top + 2\theta\,\mathrm{Id} = \lambda\Phi - \lambda\widetilde{\Phi} - \gamma\widetilde{\Psi} + 2\theta\,\mathrm{Id} + \gamma\Psi$$
$$= \Gamma_\theta - \widetilde{\Gamma} + \gamma\widetilde{\Psi}^\top.$$

It follows that the latter matrix is invertible. An application of Lemma 3.13 therefore yields that $\boldsymbol{w}$ is proportional to the vector

$$\Big(\Gamma_\theta - \widetilde{\Gamma} + \gamma\widetilde{\Psi}^\top\Big)^{-1}\mathbf{1} = A^{-1}\Big(\gamma\widetilde{\Psi}^\top + \frac{\gamma}{2}\,\mathrm{Id}\Big)^{-1}\mathbf{1} = \frac{1}{\gamma}A^{-1}(0,0,\dots,0,1)^\top,$$

where we have used the formula for $(\widetilde{\Psi}^\top + \frac{1}{2}\,\mathrm{Id})^{-1}$ given in the beginning of this proof. Since A is a nonsingular M-matrix, and thus inverse-positive by Theorem 3.8 (c), we conclude that $A^{-1}(0,0,\dots,0,1)^\top$, and in turn $\boldsymbol{w}$, have only nonnegative components. $\square$

Proof of (b)$\Rightarrow$(c) *in Theorem 2.8.* We assume $N = 1$. By definition, $\boldsymbol{w}$ is proportional to the vector

$$\det(\Gamma_\theta - \widetilde{\Gamma})(\Gamma_\theta - \widetilde{\Gamma})^{-1}\mathbf{1} = \begin{pmatrix} -0.5\gamma + 2\theta + (0.5-a)\lambda \\ 0.5(\gamma+\lambda) + 2\theta \end{pmatrix}.$$

Clearly, the second component of this vector is positive for all $a \in (0,1)$ and $\theta \geq 0$. By sending $a \uparrow 1$ one sees, however, that the first component is negative for $\theta < \theta^*$ and a sufficiently close to 1. Thus, we cannot have $\boldsymbol{w} \geq 0$ in this case. $\square$

4 Conclusion and outlook

We have considered a Nash equilibrium for two competing agents in a market impact model with general transient price impact. We have seen that without transaction costs both agents engage in a "hot-potato game", which has some similarities to certain events during the Flash Crash that have been reported in [10, 15]. We have then analyzed the behavior of equilibrium strategies as functions of transaction costs, θ, and trading frequency, N. In Theorem 2.8 we have determined the critical value of transaction costs at which the equilibrium strategies $\boldsymbol{v}$ and $\boldsymbol{v}$ become buy-only or sell-only. In Section 2.4, numerical simulations have shown that expected costs can be increasing in the trading frequency for small θ, while they generally decrease for sufficiently large θ. We have also seen that the expected costs of both agents can be lower with additional transaction costs than without. These observations provide some support for the common claim that additional transaction costs (such as a small tax or an extra spread) can, at least under certain circumstances such as during a fire sale, have a calming effect on financial markets.

In future work, we will try to provide mathematical proofs for the numerical observations in Section 2.4. These proofs will be based on a more detailed analysis of the high-frequency limit discussed in Section 2.5.

Acknowledgement: The authors thank Ria Grindel, Elias Strehle for comments that helped to improve a previous version of the manuscript.

References

[1] A. Alfonsi, A. Fruth, and A. Schied. Constrained portfolio liquidation in a limit order book model. *Banach Center Publications*, 83:9–25, 2008.

[2] A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quant. Finance*, 10:143–157, 2010.

[3] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM J. Financial Math.*, 3:511–533, 2012.

[4] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–39, 2000.

[5] A. Berman and R. J. Plemmons. *Nonnegative matrices in the mathematical sciences*, volume 9 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original.

[6] D. Bertsimas and A. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1:1–50, 1998.

[7] J.-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quant. Finance*, 4:176–190, 2004.

[8] M. K. Brunnermeier and L. H. Pedersen. Predatory trading. *Journal of Finance*, 60(4):1825–1863, August 2005.

[9] B. I. Carlin, M. S. Lobo, and S. Viswanathan. Episodic liquidity crises: cooperative and predatory trading. *Journal of Finance*, 65:2235–2274, 2007.

[10] CFTC-SEC. Findings regarding the market events of May 6, 2010. Report, 2010.

[11] D. Easley, M. López de Prado, and M. O'Hara. The microstructure of the flash crash: Flow toxicity, liquidity crashes and the probability of informed trading. *Journal of Portfolio Management*, 37(2):118–128, 2011.

[12] J. Gatheral. No-dynamic-arbitrage and market impact. *Quant. Finance*, 10:749–759, 2010.

[13] J. Gatheral and A. Schied. Dynamical models of market impact and algorithms for order execution. In J.-P. Fouque and J. Langsam, editors, *Handbook on Systemic Risk*, pages 579–602. Cambridge University Press, 2013.

[14] G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275, 07 2004.

[15] A. A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. *Preprint, available at SSRN 1686004*, 2010.

[16] C.-A. Lehalle. Market microstructure knowledge needed to control an intra-day trading process. In J.-P. Fouque and J. Langsam, editors, *Handbook on Systemic Risk*, pages 549–578. Cambridge University Press, 2013.

[17] C. C. Moallemi, B. Park, and B. Van Roy. Strategic execution in the presence of an uninformed arbitrageur. *Journal of Financial Markets*, 15(4):361 – 391, 2012.

[18] A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16:1–32, 2013.

[19] S. Predoiu, G. Shaikhet, and S. Shreve. Optimal execution in a general one-sided limit-order book. *SIAM J. Financial Math.*, 2:183–212, 2011.

[20] A. Schied and T. Zhang. A state-constrained differential game arising in optimal portfolio liquidation. *To appear in Mathematical Finance.*

[21] T. Schöneborn. Trade execution in illiquid markets. Optimal stochastic control and multi-agent equilibria. Doctoral dissertation, TU Berlin, 2008.

[22] T. Schöneborn and A. Schied. Liquidation in the face of adversity: stealth vs. sunshine trading. SSRN Preprint 1007014, 2009.