

Randomization in Optimal Execution Games

Steven Campbell* Marcel Nutz†

March 13, 2025

Abstract

We study optimal execution in markets with transient price impact in a competitive setting with N traders. Motivated by prior negative results on the existence of pure Nash equilibria, we consider randomized strategies for the traders and whether allowing such strategies can restore the existence of equilibria. We show that given a randomized strategy, there is a non-randomized strategy with strictly lower expected execution cost, and moreover this de-randomization can be achieved by a simple averaging procedure. As a consequence, Nash equilibria cannot contain randomized strategies, and non-existence of pure equilibria implies non-existence of randomized equilibria. Separately, we also establish uniqueness of equilibria. Both results hold in a general transaction cost model given by a strictly positive definite impact decay kernel and a convex trading cost.

Keywords Randomization; Optimal Execution; Transient Price Impact; N -Player Game

AMS 2020 Subject Classification 91A06; 91A15; 91G10

JEL Classification G24; C62

1 Introduction

We study optimal execution in a market with transient price impact, meaning that trades dislocate the security's price in the market but this impact diminishes over time according to a decay kernel G (see [5, 12, 26] for background and references). More specifically we are interested in a competitive setting where N traders transact in the same security and are mutually aware of their competitors' trade intention. Early works in this setting include [4, 20, 23, 24] whereas recent works with transient price impact include [1, 3, 9, 18, 21, 22, 25]. To the best of our knowledge, there are no prior works focusing on randomization in optimal execution games.

The starting point of our investigation is a surprising result of [3, 21] which use the popular Obizhaeva–Wang model [19] for price impact; i.e., the decay kernel G is exponential. In a straightforward generalization of the standard single-player optimal execution task of unwinding a given inventory in a martingale asset over a time interval $[0, T]$, it is shown

*Dept. of Statistics, Columbia University, sc5314@columbia.edu.

†Depts. of Statistics and Mathematics, Columbia University, mnutz@columbia.edu. Research supported by NSF Grants DMS-2407074, DMS-2106056.

that the N -player game does not admit a Nash equilibrium for any $N \geq 2$.¹ In view of the non-existence result, several researchers have asked if allowing for randomized execution strategies would restore existence of a Nash equilibrium. Here randomization means that traders' strategies can depend on privately observed randomization devices. This question is well motivated, as in many other games without a (pure) Nash equilibrium, existence is indeed restored by allowing for randomization. We will show, however, that this is not the case here.

More broadly, we analyze randomized strategies in an N -player game with a general impact decay kernel G and general convex trading costs C (see Section 2 for details on the setup). After mathematically formalizing a model with idiosyncratic randomization, we show that randomized strategies are not desirable in our context. Specifically, the main result shows that if some trader i considers a randomized strategy X^i , then while keeping the strategies of the competing traders fixed, replacing X^i by a de-randomized version of X^i yields a strict improvement in execution cost. In particular, this implies that *a Nash equilibrium cannot contain randomized strategies*. Mathematically, the result is driven by the strict positive definiteness of the kernel G . Our analysis further clarifies that this de-randomization can be achieved by the predictable projection onto the market filtration (which does not include the randomization devices). Concretely, this projection simply amounts to averaging the randomized strategy over the scenarios of the randomization device—essentially, taking conditional expectation given the market information.

Separately, we establish the *uniqueness of Nash equilibria*² in our general setting. This generalizes earlier results (of [21] for exponential decay kernel, $N = 2$ traders and quadratic trading costs C , and [3, 25] for the same model as [21] but general N) to a general decay kernel, general convex trading costs, and possibly randomized strategies. Like the result on de-randomization, the uniqueness proof also rests on the strict positive definiteness of the kernel.

The remainder of this note is organized as follows. Section 2 details the setup and notation, focusing on bounded decay kernels. Section 3 recalls the predictable and dual predictable projection operators from the theory of stochastic processes. Section 4 contains the main results on de-randomization of randomized strategies and the implications for Nash equilibria, while Section 5 states the uniqueness of Nash equilibria. Section 6 extends all results to the case of unbounded (weakly singular) decay kernels. Section 7 concludes with remarks about modeling choices. All proofs are reported in Section 8.

2 Setup

We consider a market where $N \geq 1$ agents trade in a single risky asset.³ We index the traders by $i \in \{1, \dots, N\}$ and denote their inventory processes by $X^i = (X_t^i)_{t \geq 0}$, where X_t^i

¹In [3], it is discussed how existence is restored when an additional quadratic cost on the trading rate (as in [10, 14]) is levied, whereas [21] discretizes time.

²Meaning that there is *at most* one equilibrium. As mentioned above, existence can fail depending on the choice of trading costs.

³The case $N = 1$ corresponds to the single-player problem where an equilibrium is just an optimal strategy. While some of our results simplify dramatically in that case, all of them remain valid.

indicates the number of shares held by trader i at time t . Each trader i is endowed with initial holdings $X_{0-}^i = x^i \in \mathbb{R}$ and trades up to the common terminal time $T > 0$. Our setup will include problems where full liquidation at time T is enforced as well as problems where terminal inventory is merely penalized.

2.1 Randomized Strategies

Our first goal is to formalize (idiosyncratically) randomized strategies. Fix a filtered probability space $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0 = (\mathcal{F}_t^0), \mathbb{P}^0)$ satisfying the usual conditions. To model randomization, we introduce auxiliary probability spaces $(\tilde{\Omega}_i, \tilde{\mathcal{F}}_i, \tilde{\mathbb{P}}_i)$, $i = 1 \dots, N$ and denote by $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ the completion of their product $\otimes_{i=1}^N (\tilde{\Omega}_i, \tilde{\mathcal{F}}_i, \tilde{\mathbb{P}}_i)$. We shall work on the space $(\Omega, \mathcal{F}, \mathbb{P})$ defined as the completion of the product space $(\Omega^0 \times \tilde{\Omega}, \mathcal{F}^0 \otimes \tilde{\mathcal{F}}, \mathbb{P}^0 \otimes \tilde{\mathbb{P}})$.

We define the market subfiltration $\mathbb{F}^M = (\mathcal{F}_t^M)_{t \geq 0}$ as the usual augmentation of the filtration $(\mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\}))_{t \geq 0}$.⁴ It represents the “public” information available to all traders. By contrast, agent i ’s subfiltration $\mathbb{F}^i = (\mathcal{F}_t^i)_{t \geq 0}$ is defined as the usual augmentation of $(\mathcal{F}_t^0 \otimes \sigma(\tilde{\pi}_i))_{t \geq 0}$, where $\tilde{\pi}_i : \tilde{\Omega} \rightarrow \tilde{\Omega}_i$ is the canonical projection $(\tilde{\omega}_1, \dots, \tilde{\omega}_N) \mapsto \tilde{\omega}_i$.⁵ We can think of $\tilde{\pi}_i$ as a random number generator whose realization can be an input of agent i ’s strategy.

Notation 2.1. Let U be an integrable random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. We denote by $\mathbb{E}[U]$ the expectation with respect to \mathbb{P} . Recall that $\Omega = \Omega^0 \times \tilde{\Omega}$ and $\mathbb{P} = \mathbb{P}^0 \otimes \tilde{\mathbb{P}}$. We denote by $\tilde{\mathbb{E}}[U]$ the marginal expectation with respect to the second factor,

$$\tilde{\mathbb{E}}[U](\omega^0) = \int_{\tilde{\Omega}} U(\omega^0, \tilde{\omega}) \tilde{\mathbb{P}}(d\tilde{\omega}).$$

(We emphasize that $\tilde{\mathbb{E}}[U]$ is a random variable, not a constant.) When $X = (X_t)$ is a process, we abuse notation and write $\tilde{\mathbb{E}}[X]$ for the process $t \mapsto \tilde{\mathbb{E}}[X_t]$.

Now that the probability space has been introduced, we can define randomized strategies, the randomization being reflected by the dependence on $\tilde{\omega}_i \in \tilde{\Omega}_i$.

Definition 2.2. We say that a process $X^i = (X_t^i)_{t \geq 0}$ on Ω is an *admissible* strategy for trader i if

- (i) X^i is càdlàg and \mathbb{F}^i -predictable,
- (ii) $X_{0-}^i = x^i$ and X_t^i is constant for $t \geq T$,
- (iii) the paths $t \mapsto X_t^i$ have (\mathbb{P} -essentially) bounded total variation.

We say that X is a *non-randomized* admissible strategy for trader i if it is also \mathbb{F}^M -predictable, whereas it is *strictly randomized* if that is not the case.

⁴The “usual augmentation” is the augmentation of the right-continuous version. In this case, $(\mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\}))_{t \geq 0}$ is already right-continuous, so the usual augmentation is just the augmentation by nullsets.

⁵ $\mathcal{F}_t^0 \otimes \sigma(\tilde{\pi}_i)$ is a short way of writing $\mathcal{F}_t^0 \otimes \mathcal{E}_1 \otimes \dots \otimes \mathcal{E}_{i-1} \otimes \tilde{\mathcal{F}}_i \otimes \mathcal{E}_{i+1} \otimes \dots \otimes \mathcal{E}_N$ where $\mathcal{E}_j = \sigma(\{\emptyset, \tilde{\Omega}_j\})$.

2.2 Price Impact

We assume that the unaffected asset price—the price that would obtain if the N agents did not trade—evolves according to a càdlàg square-integrable \mathbb{F}^M -martingale $P = (P_t)_{t \in [0, T]}$. Note that since the traders' randomization devices are independent of P , it remains a martingale in the (larger) private filtrations \mathbb{F}^i , $i = 1, \dots, N$. To describe the actual, “affected” price, we use a general impact decay kernel G ; see [12] for background and references. For ease of exposition, we first consider bounded kernels—thus avoiding pesky integrability issues—and defer the extension of our results to singular (unbounded) kernels to Section 6. Thus, we first impose the following assumption.

Assumption 2.3. The impact decay kernel $G : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is continuous and strictly positive definite in the sense of Bochner; that is,

$$\int_0^\infty \int_0^\infty G(|t - s|) dX_s dX_t > 0$$

whenever $X : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a càdlàg function of finite variation that is not constant.

Usually G is chosen to be non-increasing and convex. It is shown in [2, Proposition 2] that if $G : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is non-increasing, convex, and not constant, then the induced kernel is strictly positive definite. Two important examples are the exponential kernel $G(t) = \eta e^{-\lambda t}$ and the truncated power law kernel $G(t) = \eta(1 + \lambda t)^{-\gamma}$, where $\eta, \lambda, \gamma > 0$ are parameters.

Given a profile $\mathbf{X} = (X^1, \dots, X^N)$ of admissible strategies for traders $1, \dots, N$, we define the impact process $I = (I_t)_{t \geq 0}$ and the affected price $S = (S_t)_{t \geq 0}$ by

$$I_t = \int_0^t G(t - s) \sum_{i=1}^N dX_s^i, \quad S_t = P_t + I_t. \quad (2.1)$$

Here and throughout the paper, $\int_a^b := \int_{[a, b]}$, meaning that jumps of the integrator at a and b contribute to the integral, whereas $\int_a^{b-} := \int_{[a, b)}$ excludes the jump at b . Moreover, the “a.s.” qualifier is often suppressed.

We define the impact cost associated with $\mathbf{X} = (X^1, \dots, X^N)$ as follows. For trader i , the net proceeds from trading are

$$\int_0^T S_{t-} dX_t^i + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta S_t \Delta X_t^i, \quad (2.2)$$

where $\Delta X_t^i := X_t^i - X_{t-}^i$. Thus continuous trading at time t transacts at the price S_{t-} while a block trade of size ΔX_t^i has an execution price of $S_{t-} + \frac{G(0)}{2} \Delta S_t$. This means that trader i obtains the average execution price of all trades happening at t , and is equivalent to randomizing the order in which simultaneous block trades are executed across agents (see, e.g., [3] for a more detailed discussion).

Because our problem formulation allows for incomplete liquidation at time T , we must add to the execution costs the change in the marked-to-market value of the holdings over $[0, T]$. This change is

$$X_{0-}^i P_{0-} - X_T^i P_T \quad (2.3)$$

where we use the unaffected⁶ price P for inventory valuation. If the terminal inventory is constrained to be zero, this accounting amounts to adding a constant to the trader's cost, hence has no effect on their optimal strategy.

2.3 Additional Trading Costs and Penalizations

Finally, we introduce additional trading and terminal costs $C(X^i)$. Our exposition allows for general *convex* costs C mapping admissible strategies to $[0, \infty]$. The uniqueness result holds in that generality. The results on randomized strategies require an additional Jensen-type inequality (Assumption 4.3) which will be discussed in Section 4.

Examples for C include the quadratic costs often used in the literature. Specifically, quadratic (including zero) trading costs on the trading rate, quadratic or infinite costs on terminal inventory (the latter meaning that the complete liquidation is enforced), and/or quadratic costs on block trades as in [3, 21].

Example 2.4. Denoting by \dot{X}_t^i the time derivative of X_t^i (if it exists), let

$$C(X^i) := \frac{\varepsilon}{2} \int_0^T (\dot{X}_t^i)^2 dt + \frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t^i)^2 + \frac{\varphi}{2} (X_T^i)^2, \quad (2.4)$$

where $\varepsilon, \vartheta_t \in [0, \infty)$ and $\varphi \in [0, \infty]$. The block cost ϑ_t is bounded and measurable in t . If $\varepsilon > 0$, the formula is interpreted as $C(X^i) = \infty$ if \dot{X}^i does not exist, whereas for $\varepsilon = 0$ the first term is dropped and nothing is assumed about differentiability. This example includes quadratic (or zero) costs on the trading rate, quadratic (or zero) costs on block trades, quadratic costs on terminal inventory, or a full liquidation constraint at T (for $\varphi = \infty$).

It is clear that quadratic costs C as in Example 2.4 are convex. We will verify in Lemma 4.4 that the Jensen-type inequality of Assumption 4.3 is also satisfied. A slightly different class, also satisfying our assumptions, are discrete-time models as in [21, 25]. These can be embedded by a cost C that charges block costs as in Example 2.4 on a given grid $\mathbb{T} := \{0 = t_0 < t_1 < \dots < t_n = T\}$ of trading dates but is infinite for any strategy acting outside the grid (more precisely, the support of its total variation measure is not contained in \mathbb{T}).

2.4 Objective Function and Nash Equilibrium

Combining the additional cost C with the impact cost (2.2) and the value (2.3) of the terminal inventory, trader i has the following objective function if we fix the actions $\mathbf{X}^{-i} = (X^1, \dots, X^{i-1}, X^{i+1}, \dots, X^N)$ of the other players.

Definition 2.5. Given a profile $\mathbf{X} = (X^1, \dots, X^N)$ of admissible strategies, the objective of trader i is to minimize

$$J(X^i; \mathbf{X}^{-i}) = \mathbb{E} \left[\int_0^T S_{t-} dX_t^i + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta S_t \Delta X_t^i + (X_{0-}^i P_{0-} - X_T P_T) + C(X^i) \right]. \quad (2.5)$$

⁶This is standard in the literature; see, e.g., [17, 18]. One reason is to ensure that the objective function is convex, which can fail if the reference price is instead taken to be S_T .

The next proposition uses the martingale property of the unaffected price and provides a more explicit formula for the objective function.

Proposition 2.6. *Given a profile $\mathbf{X} = (X^1, \dots, X^N)$ of admissible strategies, the objective function $J(X^i; \mathbf{X}^{-i})$ can be expressed as*

$$J(X^i; \mathbf{X}^{-i}) = \mathbb{E} \left[\int_0^T I_{t-} dX_t^i + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta I_t \Delta X_t^i + C(X^i) \right] \quad (2.6)$$

$$= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j dX_t^i \right. \\ \left. + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta X_t^j \Delta X_t^i + C(X^i) \right]. \quad (2.7)$$

In particular, $J(X^i; \mathbf{X}^{-i}) < \infty$ as soon as $\mathbb{E}[C(X^i)] < \infty$.

Finally, the definition of Nash equilibrium is the usual one.

Definition 2.7. A profile $\mathbf{X} = (X^1, \dots, X^N)$ of admissible strategies is a *Nash equilibrium* if $J(X^i; \mathbf{X}^{-i}) < \infty$ and

$$J(Z^i; \mathbf{X}^{-i}) \geq J(X^i; \mathbf{X}^{-i})$$

for all admissible strategies Z^i of trader i , for all $i = 1, \dots, N$.

3 Preliminaries on Projections

The purpose of this section is to recall the properties of the predictable and dual predictable projection operators, and to identify them explicitly in our particular context. The takeaway is that when projecting an admissible strategy of some agent onto the market filtration or the filtration of another agent, the projection is obtained by integrating out agent i 's idiosyncratic randomization. Formally, this corresponds to the rule $\mathbb{E}[f(X, Y)|X] = \mathbb{E}[f(x, Y)]|_{x=X}$ for computing the conditional expectation when X, Y are independent random variables. For background on predictable and dual predictable projections, see for instance [15, Sections I.2d and I.3b] or [8, Section VI.2]. The following two theorems recall the basic properties.

Theorem 3.1 (predictable projection). *Let Z be a bounded $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ -measurable process and let $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions. There exists a unique (up to evanescence) \mathbb{G} -predictable process ${}^{\mathbb{P}^{\mathbb{G}}}Z$, called the predictable projection of Z , such that*

$$\mathbb{1}_{\{\tau < \infty\}} {}^{\mathbb{P}^{\mathbb{G}}}Z_\tau = \mathbb{E}[\mathbb{1}_{\{\tau < \infty\}} Z_\tau | \mathcal{G}_{\tau-}]$$

a.s. for every \mathbb{G} -predictable stopping time τ .

Theorem 3.2 (dual predictable projection). *Let Z be a $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$ -measurable process of bounded variation⁷ and let $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ be a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions. There exists a unique (up to evanescence) \mathbb{G} -predictable process $Z^{\mathbb{P}^{\mathbb{G}}}$ with bounded variation, called the dual predictable projection of Z , such that*

$$\mathbb{E} \left[\int_0^\infty \xi_s dZ_s^{\mathbb{P}^{\mathbb{G}}} \right] = \mathbb{E} \left[\int_0^\infty \xi_s dZ_s \right]$$

for all bounded \mathbb{G} -predictable processes ξ . Moreover,

$$\mathbb{E} \left[\int_0^\infty \xi_s dZ_s^{\mathbb{P}^{\mathbb{G}}} \right] = \mathbb{E} \left[\int_0^\infty \mathbb{P}^{\mathbb{G}} \xi_s dZ_s \right] = \mathbb{E} \left[\int_0^\infty \mathbb{P}^{\mathbb{G}} \xi_s dZ_s^{\mathbb{P}^{\mathbb{G}}} \right] \quad (3.1)$$

for all bounded measurable processes ξ .

Notation 3.3. For brevity, the (dual) predictable projection onto the market filtration \mathbb{F}^M will be denoted by \mathbb{p} , whereas the (dual) predictable projection onto trader i 's filtration \mathbb{F}^i will be denoted by \mathbb{p}^i ,

$$\mathbb{p} := \mathbb{p}^{\mathbb{F}^M}, \quad \mathbb{p}^i := \mathbb{p}^{\mathbb{F}^i}.$$

The next two propositions characterize the projections of an admissible strategy of some agent to (a) the market filtration \mathbb{F}^M and (b) the filtration of another agent. In either case, projecting amounts to integrating out the idiosyncratic randomization (see Notation 2.1).

Proposition 3.4. *Fix $i \in \{1, \dots, N\}$ and let Z be a càdlàg \mathbb{F}^i -predictable process of bounded variation. Then, up to evanescence,*

$$\mathbb{p}Z = Z^{\mathbb{P}} = \tilde{\mathbb{E}}[Z].^8$$

Proposition 3.5. *Fix $i, j \in \{1, \dots, N\}$ with $i \neq j$ and let Z be a càdlàg \mathbb{F}^j -predictable process of bounded variation. Then, up to evanescence,*

$$\mathbb{p}^i Z = \mathbb{p}Z = \tilde{\mathbb{E}}[Z].$$

4 Non-Existence of Randomized Equilibria

Proposition 3.4 shows that the projection of a (possibly) randomized admissible strategy X^i onto the market filtration is given by $\tilde{\mathbb{E}}[X^i]$. We record that this de-randomized strategy is again admissible.

Lemma 4.1. *Let X^i be an admissible strategy for player i . Then $\tilde{\mathbb{E}}[X^i]$ is a non-randomized admissible strategy for player i .*

⁷Since we are working with a non-trivial initial σ -field \mathcal{F}_{0-} , let us specify that bounded variation processes are understood to have bounded total variation *and* bounded initial value.

⁸In fact, the bounded variation is used only for the dual predictable projection. The identity $\mathbb{p}Z = \tilde{\mathbb{E}}[Z]$ holds for arbitrary bounded \mathbb{F}^i -predictable processes Z .

Our first proposition shows that an agent is invariant between facing (possibly) randomized strategies of the competitors or their de-randomized projections. More formally, when fixing the strategy profile \mathbf{X}^{-i} of the competitors, the expected performance of any strategy X^i against \mathbf{X}^{-i} is the same as its performance against the projection $\tilde{\mathbb{E}}[\mathbf{X}^{-i}]$ of their strategies onto the market information.

Proposition 4.2. *For any admissible strategy profile \mathbf{X} and any $i \in \{1, \dots, N\}$, the objective function J satisfies*

$$J(X^i; \mathbf{X}^{-i}) = J(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}]).$$

Our main result below uses the following condition on the convex cost C , essentially a Jensen's inequality for marginal expectations. (Note that Jensen's inequality does not always hold in infinite dimensional spaces.)

Assumption 4.3. If X is an admissible strategy for some trader i , then

$$\mathbb{E}[C(X)] \geq \mathbb{E}\left[C\left(\tilde{\mathbb{E}}[X]\right)\right].$$

Next, we verify that Assumption 4.3 is satisfied for the quadratic costs that are predominantly used in the literature.

Lemma 4.4. *Assumption 4.3 holds for the quadratic costs C specified in Example 2.4.*

We can now state our first main result. It shows that de-randomizing a randomized strategy by projecting leads to an improved performance while holding competitors' strategies fixed. In particular, this implies the non-existence of (strictly) randomized equilibria.

Theorem 4.5. *Let C satisfy Assumption 4.3 and let $\mathbf{X} = (X^1, \dots, X^N)$ be an admissible strategy profile where X^i is strictly randomized. Then $Z^i := \tilde{\mathbb{E}}[X^i]$ is a non-randomized admissible strategy for trader i and*

$$J(X^i; \mathbf{X}^{-i}) > J(Z^i; \mathbf{X}^{-i}).$$

That is, a strictly randomized strategy X^i can be strictly improved by unilaterally deviating to the non-randomized strategy $Z^i = \tilde{\mathbb{E}}[X^i]$. In particular, a Nash equilibrium profile cannot contain strictly randomized strategies.

The proof shows that this result is driven by the strict positive definiteness of the impact decay kernel. The assumption on C , which is a non-strict inequality, merely ensures that C does not counteract that property of the kernel.

5 Uniqueness of Nash Equilibria

We show the uniqueness of Nash equilibria in our general framework.⁹ The first result in this direction is due to [21], where uniqueness is shown for the exponential decay kernel (i.e.,

⁹To be clear, Assumption 4.3 is not in force in this section.

the Obizhaeva–Wang model), $N = 2$ traders and a quadratic cost C that is a special case of Example 2.4. This result was extended to similar games with an arbitrary finite number N of traders in [3, 25]. A different uniqueness argument, based on Fredholm equations of the second kind, was given in [1] for linear-quadratic models that are regularized by a quadratic cost on the trading rate. Here, we establish a general version of the uniqueness property including general convex costs C , a general decay kernel, and (possibly) randomized strategies. We emphasize that the theorem merely asserts uniqueness; existence need not hold under the present assumptions (see, e.g., [3]).

Theorem 5.1. *There is at most one Nash equilibrium.*

Following the initial argument of [21], the proof in Section 8.5 rests on the strict positive definiteness of the decay kernel. The cost C is assumed to be convex but not necessarily strictly convex ($C \equiv 0$ is allowed). Again, the convexity assumption makes sure that the cost does not counteract the properties of the kernel, but on its own clearly does not yield uniqueness.

6 Extension to Singular Kernels

In this section we extend our results to impact decay kernels with $G(0+) = \infty$. See [12] for the significance of such kernels. The most important example is the power-law kernel $G(t) = t^\gamma$ where $\gamma \in (0, 1)$; cf. [11]. We put ourselves in the framework of [13] for (weakly) singular kernels and impose the following throughout this section.

Assumption 6.1 (singular kernel). The function $G : \mathbb{R}_+ \rightarrow [0, \infty]$ is convex and non-increasing, continuous and finite on $(0, \infty)$ with $\int_0^1 G(t)dt < \infty$, and $G(0) = \lim_{t \downarrow 0} G(t) = \infty$.

Note that, for convenience of exposition, we have made $G(0) = \infty$ part of the assumption: we focus on the singular case, rather than including the bounded case of Assumption 6.1 whose results were already stated. The definition of admissibility needs to be amended as follows.

Definition 6.2. A process $X^i = (X_t^i)_{t \geq 0}$ is an *admissible* strategy for trader i if it satisfies Definition 2.2 and in addition

$$\mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|X^i|_s d|X^i|_t \right] < \infty. \quad (6.1)$$

While (6.1) was automatic for bounded G (as X^i is of bounded variation), this additional condition is needed to ensure that the quantities of interest are well-defined when G is unbounded. The next remark emphasizes that in the present setting with $G(0) = \infty$, *admissible strategies cannot have jumps*; that is, there are no block trades.

Remark 6.3. The integral $\int_0^T \int_0^T G(|t-s|) d|X|_s d|X|_t$ is necessarily infinite if X has a jump, as $\int_0^T \int_0^T G(|t-s|) d|X|_s d|X|_t \geq G(0) |\Delta X_t|^2$ for any t and $G(0) = \infty$. Thus, in view of (6.1), admissible strategies cannot have jumps in the present setting. In financial terms, block trades have infinite price impact and hence infinite execution cost.

As in the previous sections, our main results will be driven by the strict positive definiteness of the kernel. Let us record that strict positive definiteness indeed holds. As detailed in the proof, this property can be seen from a Fourier-type representation shown in [13].

Lemma 6.4. *Let G satisfy Assumption 6.1. Then G is strictly positive definite; i.e.,*

$$\int_0^T \int_0^T G(|t-s|) dX_s dX_t > 0$$

whenever $X : [0, T] \rightarrow \mathbb{R}$ is a càdlàg function of finite variation that is not constant and satisfies $\int_0^T \int_0^T G(|t-s|) d|X|_s d|X|_t < \infty$.

The following approximation will enable us to exploit our results for bounded kernels.

Remark 6.5 (monotone approximation). Let G satisfy Assumption 6.1. Then given any $n > 0$ such that G is not constant on $[1/n, \infty)$, the function $G^n(t) := G(t + 1/n)$, $t \in \mathbb{R}_+$ is a bounded kernel satisfying Assumption 2.3 (by the sufficient condition mentioned below Assumption 2.3). Moreover, we have the monotone limit $G^n(t) \uparrow G(t)$ as $n \uparrow \infty$. This will be useful for monotone convergence arguments.

The objective function J is the same as in Section 2, with the following caveats. First, the expressions involving jump terms can be dropped by Remark 6.3. Second, in contrast to the case of bounded kernels, it may not be obvious that various integrals are well defined, and some steps in the proof of Proposition 2.6 need additional explanations. We provide those in Section 8.6, and we also verify by elementary arguments that (6.1) and positive definiteness imply that all expressions are well defined. We thus have the following simplified expression for Proposition 2.6.

Proposition 6.6. *Given a profile $\mathbf{X} = (X^1, \dots, X^N)$ of admissible strategies, the objective function $J(X^i; \mathbf{X}^{-i})$ can be expressed as*

$$\begin{aligned} J(X^i; \mathbf{X}^{-i}) &= \mathbb{E} \left[\int_0^T I_{t-} dX_t^i + C(X^i) \right] \\ &= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j dX_t^i + C(X^i) \right]. \end{aligned} \tag{6.2}$$

Moreover, $|J(X^i; \mathbf{X}^{-i})| < \infty$ as soon as $\mathbb{E}[C(X^i)] < \infty$.

A technical difficulty in the present setup is that the impact process I_t may be infinite for some exceptional set of t , even when strategies are admissible. This does not contradict the preceding statements because the objective function only includes certain integrals of the impact process, not the impact process at a fixed time. Thus, strictly speaking, there is no need to address the difficulty. Let us mention, however, the following result of [13, Proposition 2.27].

Remark 6.7. Let $X : [0, T] \rightarrow \mathbb{R}$ be càdlàg function of finite variation with $\int_0^T \int_0^T G(|t - s|) d|X|_s d|X|_t < \infty$. Then $\int_0^t G(t - s) dX_s$ is well defined and finite for *quasi* every $t \in [0, T]$, meaning that the set of exceptional t is ν -null for every finite Borel measure ν satisfying $\iint G(|t - s|) \nu(ds) \nu(dt) < \infty$.

We may apply this with X being a path of an admissible strategy and μ being the total variation measure of the path of another admissible strategy. But, as mentioned, the elementary results in Section 8.6 suffice for our purpose.

We can now formally state that all our results extend to the singular setting.

Theorem 6.8. *Let G satisfy Assumption 6.1 and define admissibility as in Definition 6.2. Then Lemma 4.1, Proposition 4.2, Theorem 4.5 and Theorem 5.1 hold as stated.*

7 Conclusion

In our model of an optimal execution game with transient price impact, we have shown that Nash equilibria are unique and do not contain randomized strategies. Our starting point was the non-existence of pure Nash equilibria in certain optimal execution games with full information, and the question whether allowing for randomization restores existence. We have chosen a direct extension of the full information setup predominant in the literature. While in some other games, such an extension enables Nash equilibria, that is not the case here. Our setting is the strongest form of randomization because traders cannot see the impact of each other's actions, thus gives a conservative answer to the initial question.

The unaffected price is treated as known in the full information model, but ultimately it does not enter the objective (Proposition 2.6) and hence the strategies do not depend on it, as long as the martingale assumption holds. On the other hand, the randomization implies that the affected price is not measurable in agents' filtrations. Let us discuss briefly another possible model, where agents' information sets additionally include the affected price (or, the impact process, which is equivalent since the unaffected price is known). Then, the objective remains in terms of observable quantities (Definition 2.5) and the problem reduces to the non-randomized full information setting (i.e., a market with a common filtration for all traders). It is then unsurprising that this randomization does not prevent non-existence of equilibria.

We emphasize that, like the previous literature, we start with a full information setup where the unaffected price is known to agents. In the financial reality, the affected price is observed but the unaffected price is not. Under the assumption of a martingale unaffected price, but without observation of the latter, observing the affected price allows agents to filter for the impact process. In such a setup, agents have partial information about their competitors' strategies even in the presence of randomization. In the (realistic) regime where the unaffected price is quite noisy and price impact is fairly small relative to that noise, this inference would be quite limited.¹⁰ Our setup, where the affected price is not measurable in agents' filtrations, roughly corresponds to the limiting regime of high noise where no

¹⁰In reality, traders might gain additional information from various sources ranging from observing fills of their own orders in various venues to trade volume data.

information is gained by adding the observation of the affected price. We hope to address the partial information setup rigorously in future work. Results in that direction are scarce; see [16] for a model where all impact is permanent and [6, 7] for a mean field model.

8 Proofs

8.1 Notation

Recall that we write $\int_a^b H_t dX_t$ to denote the integral over $[a, b] \subset \mathbb{R}$ for suitable integrands H and integrators X . In particular, $\int_a^b H_t dX_t$ includes the contribution $H_a \Delta X_a$ of a jump in X at $t = a$ and similarly at $t = b$. Whereas, $\int_a^{b-} H_t dX_t$ is the integral over $[a, b)$, excluding the jump at b .

The subsequent subsections, except for the last one, all consider a (bounded) kernel G satisfying Assumption 2.3. The last subsection deals with the extension to singular kernels, meaning that G satisfies Assumption 6.1.

8.2 Proofs for Section 2 (Objective Function)

Proof of Proposition 2.6. Expanding (2.5) gives

$$J(X^i; \mathbf{X}^{-i}) = \mathbb{E} \left[\int_0^T P_{t-} dX_t^i + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta P_t \Delta X_t^i + \int_0^T I_{t-} dX_t^i + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta I_t \Delta X_t^i + (X_{0-}^i P_{0-} - X_T^i P_T) + C(X^i) \right].$$

Using integration by parts,

$$\int_0^T P_{t-} dX_t^i = P_T X_T^i - P_{0-} X_{0-}^i - \int_0^T X_{t-}^i dP_t - [X^i, P]_T$$

and thus

$$J(X^i; \mathbf{X}^{-i}) = \mathbb{E} \left[- \int_0^T X_{t-}^i dP_t - [X^i, P]_T + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta P_t \Delta X_t^i + \int_0^T I_{t-} dX_t^i + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta I_t \Delta X_t^i + C(X^i) \right].$$

As P is a square-integrable $(\mathbb{F}^i, \mathbb{P})$ -martingale and X^i is a bounded predictable process, both $\int_0^t X_s^i dP_s$ and $[X^i, P]_t = \int_0^t \Delta X_s^i dP_s$ are true $(\mathbb{F}^i, \mathbb{P})$ -martingales, hence have vanishing expectation. Moreover,

$$\frac{G(0)}{2} \sum_{t \in [0, T]} \Delta X_t^i \Delta P_t = \frac{G(0)}{2} [X^i, P]_T,$$

so its expectation also vanishes. This proves the claimed representation (2.6).

For the remaining representation (2.7), we insert the definition (2.1) of I into (2.6) to get

$$J(X^i; \mathbf{X}^{-i}) = \mathbb{E} \left[C(X^i) + \int_0^T \int_0^{t-} G(t-s) \sum_{j=1}^N dX_s^j dX_t^i + \frac{G(0)}{2} \sum_{j=1}^N \sum_{t \in [0, T]} \Delta X_t^j \Delta X_t^i \right]. \quad (8.1)$$

Splitting the integral and applying Fubini's theorem to interchange the order of integration,

$$\begin{aligned} & \int_0^T \int_0^{t-} G(t-s) dX_s^i dX_t^i \\ &= \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dX_s^i dX_t^i + \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dX_s^i dX_t^i \\ &= \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dX_s^i dX_t^i + \frac{1}{2} \int_0^T \int_s^T G(t-s) dX_t^i dX_s^i - \frac{G(0)}{2} \sum_{s \in [0, T]} (\Delta X_s^i)^2 \\ &= \frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i - \frac{G(0)}{2} \sum_{t \in [0, T]} (\Delta X_t^i)^2. \end{aligned}$$

Substituting this expression into (8.1) completes the proof of (2.6). Finiteness of the objective when $\mathbb{E}[C(X^i)] < \infty$ is clear from (2.6). \square

8.3 Proofs for Section 3 (Projections)

Lemma 8.1. *The filtration $(\mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\}))_{t \geq 0}$ is right continuous and hence its augmentation is the market filtration \mathbb{F}^M .*

Proof. It is elementary to verify the right-continuity. The second claim follows as \mathbb{F}^M was defined as the usual augmentation of $(\mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\}))_{t \geq 0}$. \square

We record the following consequence of Lemma 8.1 for ease of reference.

Lemma 8.2. *Let $t \geq 0$ and let U be an integrable \mathcal{F} -measurable random variable. Then $\mathbb{E} [U | \mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\})] = \mathbb{E} [U | \mathcal{F}_t^M]$ \mathbb{P} -a.s.*

The next lemma formalizes that conditioning a random variable from agent i 's filtration to the market filtration is equivalent to integrating out the randomization.

Lemma 8.3. *Let $t \geq 0$ and let U be an integrable \mathcal{F}_t^i -measurable random variable for some $i \in \{1, \dots, N\}$. Then $\tilde{\mathbb{E}}[U] = \mathbb{E}[U | \mathcal{F}_t^M]$ \mathbb{P} -a.s.*

Proof. By Lemma 8.2 it suffices to show that $\tilde{\mathbb{E}}[U]$ is a version of $\mathbb{E}[U | \mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\})]$. The measurability of marginal expectations on completed product spaces entails that $\omega^0 \mapsto \tilde{\mathbb{E}}[U](\omega^0)$ is \mathcal{F}_t^0 -measurable, and hence that $(\omega^0, \tilde{\omega}) \mapsto \tilde{\mathbb{E}}[U](\omega^0, \tilde{\omega})$ is $\mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\})$ -measurable. Let $A \in \mathcal{F}_t^0 \otimes \sigma(\{\emptyset, \tilde{\Omega}\})$, so that $A = A' \times \tilde{\Omega}$ for some $A' \in \mathcal{F}_t^0$. By the definition of the product measure \mathbb{P} , we have $\int_A U d\mathbb{P} = \int_{A'} \int_{\tilde{\Omega}} U d\tilde{\mathbb{P}} d\mathbb{P}^0 = \int_{A'} \tilde{\mathbb{E}}[U] d\mathbb{P}^0 = \int_A \tilde{\mathbb{E}}[U] d\mathbb{P}$. \square

We can now detail the proof of Proposition 3.4, which is a version of Lemma 8.3 for processes.

Proof of Proposition 3.4. If Z is an elementary \mathbb{F}^i -predictable process then it takes the form

$$Z = Z_{0-}\mathbf{1}_{0-}(\cdot) + Z_0\mathbf{1}_0(\cdot) + \sum_{k=1}^d Z_{s_k}\mathbf{1}_{(s_k, t_k]}(\cdot)$$

for some $s_k < t_k$, $\mathcal{F}_{s_k}^i$ -measurable Z_{s_k} , and \mathcal{F}_{0-}^i -measurable Z_{0-}, Z_0 . Taking expectations with respect to $\tilde{\mathbb{P}}$ and invoking Lemma 8.3, it is clear that $\tilde{\mathbb{E}}[Z]$ is an elementary \mathbb{F}^M -predictable process. By a monotone class argument it follows that $\tilde{\mathbb{E}}[Z]$ is \mathbb{F}^M -predictable also for a general bounded predictable processes Z .

Let Z be bounded and \mathbb{F}^i -predictable; we show that ${}^pZ = \tilde{\mathbb{E}}[Z]$. Since $\tilde{\mathbb{E}}[Z]$ is predictable, $\tilde{\mathbb{E}}[Z_\tau]$ is $\mathcal{F}_{\tau-}^M$ -measurable for all \mathbb{F}^M -predictable times τ .¹¹ Thus, we may check the definition of the conditional expectation. Recall that by definition,

$$\mathcal{F}_{\tau-}^M = \sigma\left(\{B \cap \{t < \tau\} : t \geq 0, B \in \mathcal{F}_t^M\} \cup \mathcal{F}_{0-}^M\right).$$

All of these generating sets are necessarily of the form $(A' \times \tilde{\Omega}) \cup N$ for a \mathbb{P} -null set N , and this extends to the σ -field. For any $A = (A' \times \tilde{\Omega}) \cup N \in \mathcal{F}_{\tau-}^M$, we argue as in the proof of Lemma 8.3 that $\int_A Z_\tau d\mathbb{P} = \int_A \tilde{\mathbb{E}}[Z_\tau] d\mathbb{P}$. We conclude that $\mathbb{E}[Z_\tau | \mathcal{F}_{\tau-}^M] = \tilde{\mathbb{E}}[Z_\tau]$ and so,

$$\mathbb{E}[1_{\{\tau < \infty\}} Z_\tau | \mathcal{F}_{\tau-}^M] = 1_{\{\tau < \infty\}} \tilde{\mathbb{E}}[Z_\tau].$$

By the uniqueness up to evanescence in Theorem 3.1, ${}^pZ = \tilde{\mathbb{E}}[Z]$.

We turn to the claim that $Z^p = \tilde{\mathbb{E}}[Z]$ when Z is càdlàg and of bounded variation. By the above, $\tilde{\mathbb{E}}[Z]$ is a \mathbb{F}^M -predictable process of bounded variation. In particular, $\tilde{\mathbb{E}}[Z] = \tilde{\mathbb{E}}[Z]^p$. To show that $Z^p = \tilde{\mathbb{E}}[Z]$, it then suffices (by [8, Theorem VI.75(b)]) to check that

$$\mathbb{E}[Z_0 | \mathcal{F}_{0-}^M] = \mathbb{E}[\tilde{\mathbb{E}}[Z_0] | \mathcal{F}_{0-}^M], \quad \mathbb{E}[Z_\infty - Z_t | \mathcal{F}_t^M] = \mathbb{E}[\tilde{\mathbb{E}}[Z_\infty] - \tilde{\mathbb{E}}[Z_t] | \mathcal{F}_t^M], \quad t \geq 0.$$

We prove the second part; the first is similar. Fix $t \geq 0$. To simplify notation, let $\zeta = Z_\infty - Z_t$. Moreover, note that it suffices to show

$$\mathbb{E}[\zeta | \mathcal{F}_\infty^M] = \tilde{\mathbb{E}}[\zeta], \tag{8.2}$$

as an application of the tower property then yields the claim. To show (8.2), note that by construction of \mathbb{F}^i we almost surely have $\zeta(\omega^0, \tilde{\omega}) = \bar{\zeta}(\omega^0, \tilde{\omega}_i)$ for some measurable function $\bar{\zeta}$. Using the product form of \mathbb{P} ,

$$\mathbb{E}[\zeta | \mathcal{F}_\infty^M](\omega^0, \tilde{\omega}) = \int \bar{\zeta}(\omega^0, \tilde{\omega}_i) \tilde{\mathbb{P}}_i(d\tilde{\omega}_i) = \tilde{\mathbb{E}}[\zeta](\omega^0),$$

completing the proof. □

¹¹Our notation is ambiguous, but note that the integral $\tilde{\mathbb{E}}[Z_\tau]$ indeed coincides the the evaluation of the process $t \mapsto \tilde{\mathbb{E}}[Z_t]$ at $t = \tau$ when τ is an \mathbb{F}^M -stopping time.

Finally, we prove the analogue for projecting onto the filtration of another agent.

Proof of Proposition 3.5. By Proposition 3.4, $\tilde{\mathbb{E}}[Z]$ is an \mathbb{F}^M -predictable process. In particular, $\tilde{\mathbb{E}}[Z]$ is predictable in the filtration $\mathbb{F}^i \supset \mathbb{F}^M$ and it suffices to check that

$$\mathbf{1}_{\{\tau < \infty\}} \tilde{\mathbb{E}}[Z_t] |_{t=\tau} = \mathbb{E}[\mathbf{1}_{\{\tau < \infty\}} Z_\tau | \mathcal{F}_{\tau-}^i]$$

for all \mathbb{F}^i -predictable times τ . By construction of \mathbb{F}^i and \mathbb{F}^j we almost surely have

$$\tau(\omega^0, \tilde{\omega}) = \bar{\tau}(\omega^0, \tilde{\omega}_i), \quad \mathbf{1}_{\{t < \infty\}} Z_t(\omega^0, \tilde{\omega}) = \bar{Z}(t, \omega^0, \tilde{\omega}_j)$$

for some measurable functions $\bar{\tau}, \bar{Z}$. Hence $\mathbf{1}_{\{\tau < \infty\}} Z_\tau$ is the composition

$$(\mathbf{1}_{\{\tau < \infty\}} Z_\tau)(\omega^0, \tilde{\omega}) = \bar{Z}(\bar{\tau}(\omega^0, \tilde{\omega}_i), \omega^0, \tilde{\omega}_j)$$

and by the product form of \mathbb{P} ,

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\{\tau < \infty\}} Z_\tau | \mathcal{F}_\infty^i](\omega^0, \tilde{\omega}) &= \int \bar{Z}(\bar{\tau}(\omega^0, \tilde{\omega}_i), \omega^0, \tilde{\omega}_j) \tilde{\mathbb{P}}_j(d\tilde{\omega}_j) \\ &= (\mathbf{1}_{\{\tau < \infty\}} \tilde{\mathbb{E}}[Z_t] |_{t=\tau})(\omega^0, \tilde{\omega}). \end{aligned}$$

As we already know that the right hand side is measurable with respect to $\mathcal{F}_{\tau-}^i \subset \mathcal{F}_\infty^i$, the claim follows by the tower property of conditional expectation. \square

8.4 Proofs for Section 4 (De-Randomization)

8.4.1 Admissibility of $\tilde{\mathbb{E}}[X]$ (Lemma 4.1)

In what follows we denote by $TV(X; [a, b])$ the total variation of the path $t \mapsto X_t$ on $[a, b]$. Our first result is straightforward.

Lemma 8.4. $TV(\tilde{\mathbb{E}}[X]; [a, b]) \leq \tilde{\mathbb{E}}[TV(X; [a, b])]$ for any $0 \leq a \leq b \leq \infty$

Proof. For any partition $\Pi = \{t_0, \dots, t_n\}$ of $[a, b]$,

$$\sum_{i=0}^{n-1} |\tilde{\mathbb{E}}[X_{t_{i+1}}] - \tilde{\mathbb{E}}[X_{t_i}]| \leq \tilde{\mathbb{E}} \left[\sum_{i=0}^{n-1} |X_{t_{i+1}} - X_{t_i}| \right] \leq \tilde{\mathbb{E}}[TV(X; [a, b])].$$

Taking the supremum over partitions Π yields the claim. \square

Proof of Lemma 4.1. As $\tilde{\mathbb{E}}[X^i] = {}^p X^i$ (cf. Proposition 3.4), it is non-randomized and \mathbb{F}^M -predictable, hence \mathbb{F}^i -predictable. Using Lemma 8.4, it is then straightforward to verify the conditions of Definition 2.2. \square

8.4.2 Proof of Proposition 4.2

Proof of Proposition 4.2. We first note a trivial case of (3.1): if the process Z in Theorem 3.2 is itself \mathbb{G} -predictable, clearly $Z^{\mathbb{P}^{\mathbb{G}}} = Z$, and then (3.1) includes the identity

$$\mathbb{E} \left[\int_0^\infty \xi_s dZ_s \right] = \mathbb{E} \left[\int_0^\infty \mathbb{P}^{\mathbb{G}} \xi_s dZ_s \right].$$

Consider (2.7) and more specifically the terms that depend on \mathbf{X}^{-i} . Denoting those by $J^{-i}(X^i; \mathbf{X}^{-i})$, applying Theorem 3.2, and using the \mathbb{F}^i -predictability of X^i as just noted,

$$\begin{aligned} J^{-i}(X^i; \mathbf{X}^{-i}) &:= \mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j dX_t^i + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta X_t^j \Delta X_t^i \right] \\ &= \mathbb{E} \left[\int_0^T \mathbb{P}^i \left(\int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j \right) dX_t^i + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \mathbb{P}^i(\Delta X_t^j) \Delta X_t^i \right]. \end{aligned}$$

(To apply Theorem 3.2, note that the sum is a special case of an integral, $\sum_{t \in [0, T]} \Delta X_t^j \Delta X_t^i = \int_0^T \Delta X_t^j dX_t^i$.) Applying Proposition 3.5 we then get

$$\begin{aligned} &\mathbb{E} \left[\int_0^T \mathbb{P}^i \left(\int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j \right) dX_t^i + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \mathbb{P}^i(\Delta X_t^j) \Delta X_t^i \right] \\ &= \mathbb{E} \left[\int_0^T \mathbb{P} \left(\int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j \right) dX_t^i + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \mathbb{P}(\Delta X_t^j) \Delta X_t^i \right] \\ &= \mathbb{E} \left[\int_0^T \mathbb{P} \left(\int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j \right) dX_t^i + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta((X_t^j)^{\mathbb{P}}) \Delta X_t^i \right], \quad (8.3) \end{aligned}$$

where the last equality used the general fact that $\mathbb{P}(\Delta A) = \Delta(A^{\mathbb{P}})$ for any càdlàg adapted process A with integrable variation (see [15, 3.21, p. 33]). Denote the first integrand in (8.3) by

$$I_t^{-i} := \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j.$$

From Proposition 3.4 we know that $\mathbb{P}I_t^{-i} = \tilde{\mathbb{E}}[I_t^{-i}]$ and we show below that

$$\tilde{\mathbb{E}}[I_t^{-i}] = \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j]. \quad (8.4)$$

Next, we substitute these identities into the first term of (8.3). Using also that $(X^j)^{\mathbb{P}} = \tilde{\mathbb{E}}[X^j]$ by Proposition 3.4 to rewrite the second term of (8.3), we get

$$J^{-i}(X^i; \mathbf{X}^{-i}) = \mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j] dX_t^i + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta \tilde{\mathbb{E}}[X_s^j] \Delta X_t^i \right].$$

Comparing with (2.7) we see that the right hand side equals $J^{-i}(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}])$, thus establishing that $J(X^i; \mathbf{X}^{-i}) = J(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}])$.

It remains to show (8.4). This is clear for $t = 0$ so we may focus on $t > 0$. Recall that G is continuous on $[0, T]$. Suppose for a moment that G is even continuously differentiable. Then integrating by parts gives

$$\int_0^r G(r-s) \sum_{j \neq i} dX_s^j = G(0) \sum_{j \neq i} X_r^j - G(r) \sum_{j \neq i} X_{0-}^j + \int_0^r G'(r-s) \sum_{j \neq i} X_s^j ds.$$

Taking expectations, applying Fubini's theorem, and reversing the integration by parts yields

$$\begin{aligned} \tilde{\mathbb{E}} \left[\int_0^r G(r-s) \sum_{j \neq i} dX_s^j \right] &= G(0) \sum_{j \neq i} \tilde{\mathbb{E}}[X_r^j] - G(r) \sum_{j \neq i} \tilde{\mathbb{E}}[X_{0-}^j] + \int_0^r G'(r-s) \sum_{j \neq i} \tilde{\mathbb{E}}[X_s^j] ds \\ &= \int_0^r G(r-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j]. \end{aligned} \quad (8.5)$$

In fact, these steps hold for any continuously differentiable function G . The general continuous kernel G can be approximated uniformly on $[0, T]$ with continuously differentiable functions, hence it follows that (8.5) also holds for the general kernel G . The claim (8.4) now follows by taking the limit $r \uparrow t$ and using that G is continuous and bounded. \square

8.4.3 Proof of Lemma 4.4

Next, we show that the quadratic costs of Example 2.4 satisfy the Jensen-type inequality of Assumption 4.3.

Proof of Lemma 4.4. If $\mathbb{E}[C(X)] = \infty$, the conclusion is immediate. We therefore fix an admissible strategy X for some trader i with $\mathbb{E}[C(X)] < \infty$. Observe that by Fubini's theorem,

$$\mathbb{E}[C(X)] = \int_{\Omega^0} \tilde{\mathbb{E}}[C(X)] d\mathbb{P}^0 = \int_{\Omega^0} \int_{\tilde{\Omega}} \tilde{\mathbb{E}}[C(X)] d\tilde{\mathbb{P}} d\mathbb{P}^0 = \mathbb{E} \left[\tilde{\mathbb{E}}[C(X)] \right]. \quad (8.6)$$

Invoking Fubini's theorem once again with a pointwise application of Jensen's inequality on the quadratic function $z \mapsto z^2$ for $z \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E} \left[\tilde{\mathbb{E}}[C(X)] \right] &= \mathbb{E} \left[\tilde{\mathbb{E}} \left[\frac{\varepsilon}{2} \int_0^T (\dot{X}_t)^2 dt + \frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 + \frac{\varphi}{2} (X_T)^2 \right] \right] \\ &\geq \mathbb{E} \left[\frac{\varepsilon}{2} \int_0^T \left(\tilde{\mathbb{E}}[\dot{X}_t] \right)^2 dt + \tilde{\mathbb{E}} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 \right] + \frac{\varphi}{2} \left(\tilde{\mathbb{E}}[X_T] \right)^2 \right]. \end{aligned} \quad (8.7)$$

With a view towards the first term, we next argue that if $\varepsilon > 0$, then $\frac{d}{dt} \tilde{\mathbb{E}}[X_t] = \tilde{\mathbb{E}}[\dot{X}_t] dt \otimes d\mathbb{P}$ almost surely. Indeed, if $\varepsilon > 0$, then by the definition in Example 2.4 and $\mathbb{E}[C(X)] < \infty$ we know that X is absolutely continuous. Thus, given the initial value $\tilde{\mathbb{E}}[X_0] = X_0 = x^i \in \mathbb{R}$,

$$\tilde{\mathbb{E}}[X_t] = x^i + \tilde{\mathbb{E}} \left[\int_0^t \dot{X}_s ds \right].$$

Since X is admissible, it has \mathbb{P} -essentially bounded variation and in particular $\int_0^t |\dot{X}_s| ds$ is uniformly bounded. Thus, we may apply Fubini's theorem to obtain $\tilde{\mathbb{E}}[X_t] = x^i + \int_0^t \tilde{\mathbb{E}}[\dot{X}_s] ds$. In other words, $\tilde{\mathbb{E}}[X_t]$ is absolutely continuous with derivative $\frac{d}{dt} \tilde{\mathbb{E}}[X_t] = \tilde{\mathbb{E}}[\dot{X}_t]$ $dt \otimes d\mathbb{P}$ almost surely, and we have

$$\mathbb{E} \left[\tilde{\mathbb{E}} \left[\frac{\varepsilon}{2} \int_0^T (\dot{X}_t)^2 dt \right] \right] \geq \mathbb{E} \left[\frac{\varepsilon}{2} \int_0^T \left(\frac{d}{dt} \tilde{\mathbb{E}}[X_t] \right)^2 dt \right]. \quad (8.8)$$

It remains to treat the second term in (8.7). By Theorem 3.2 and the identity $X^p = \tilde{\mathbb{E}}[X]$ from Proposition 3.4,

$$\begin{aligned} 0 &\leq \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t \left(\Delta X_t^i - \Delta \tilde{\mathbb{E}}[X_t^i] \right)^2 \right] \\ &= \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t \left((\Delta X_t)^2 - 2\Delta X_t \Delta \tilde{\mathbb{E}}[X_t] + (\Delta \tilde{\mathbb{E}}[X_t])^2 \right) \right] \\ &= \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 - \frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta \tilde{\mathbb{E}}[X_t])^2 \right] \\ &\quad + \mathbb{E} \left[\sum_{t \in [0, T]} \vartheta_t (\Delta \tilde{\mathbb{E}}[X_t])^2 - \sum_{t \in [0, T]} \vartheta_t \Delta X_t \Delta \tilde{\mathbb{E}}[X_t] \right] \\ &= \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 - \frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta \tilde{\mathbb{E}}[X_t])^2 \right] \\ &\quad + \mathbb{E} \left[\sum_{t \in [0, T]} \vartheta_t (\Delta \tilde{\mathbb{E}}[X_t])^2 - \sum_{t \in [0, T]} \vartheta_t \Delta X_t^p \Delta \tilde{\mathbb{E}}[X_t] \right] \\ &= \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 - \frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta \tilde{\mathbb{E}}[X_t])^2 \right]. \end{aligned}$$

By rearranging we conclude that

$$\mathbb{E} \left[\tilde{\mathbb{E}} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 \right] \right] = \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta X_t)^2 \right] \geq \mathbb{E} \left[\frac{1}{2} \sum_{t \in [0, T]} \vartheta_t (\Delta \tilde{\mathbb{E}}[X_t])^2 \right]. \quad (8.9)$$

Combining Eqs. (8.6) to (8.9) completes the proof. \square

8.4.4 Proof of Theorem 4.5

We first prove the following technical lemma.

Lemma 8.5. For any kernel G satisfying Assumption 2.3 and any admissible strategy X ,

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) d(X_s - \tilde{\mathbb{E}}[X_s]) d(X_t - \tilde{\mathbb{E}}[X_t]) \right] \\ &= \mathbb{E} \left[\frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) dX_s dX_t - \frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) d\tilde{\mathbb{E}}[X_s] d\tilde{\mathbb{E}}[X_t] \right]. \end{aligned}$$

Proof. Since X and $\tilde{\mathbb{E}}[X]$ are constant after T it suffices to check the equality for upper limit of integration T . We have

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) d(X_s^i - \tilde{\mathbb{E}}[X_s^i]) d(X_t^i - \tilde{\mathbb{E}}[X_t^i]) \right] \\ &= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i + \frac{1}{2} \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right. \\ &\quad \left. - \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] dX_t^i \right] \\ &= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i - \frac{1}{2} \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right. \\ &\quad \left. + \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] - \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] dX_t^i \right]. \end{aligned} \tag{8.10}$$

Next, note that for arbitrary Lebesgue–Stieltjes integrators M, N we have

$$\begin{aligned} & \int_0^T \int_0^{t-} G(t-s) dM_s dN_t \\ &= \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dM_s dN_t + \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dM_s dN_t \\ &= \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dM_s dN_t + \frac{1}{2} \int_0^T \int_s^T G(t-s) dN_t dM_s - \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta M_t \Delta N_t \\ &= \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dM_s dN_t + \frac{1}{2} \int_0^T \int_t^T G(s-t) dM_s dN_t - \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta M_t \Delta N_t \\ &\quad + \frac{1}{2} \int_0^T \int_t^T G(s-t) dN_s dM_t - \frac{1}{2} \int_0^T \int_t^T G(s-t) dM_s dN_t \\ &= \frac{1}{2} \int_0^T \int_0^T G(|t-s|) dM_s dN_t - \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta M_t \Delta N_t \\ &\quad + \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dM_s dN_t - \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dN_s dM_t. \end{aligned}$$

Rearranging, this leads to the identity

$$\begin{aligned} \frac{1}{2} \int_0^T \int_0^T G(|t-s|) dM_s dN_t &= \int_0^T \int_0^{t-} G(t-s) dM_s dN_t + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta M_t \Delta N_t \\ &\quad + \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dN_s dM_t - \frac{1}{2} \int_0^T \int_0^{t-} G(t-s) dM_s dN_t. \end{aligned} \quad (8.11)$$

In particular, if $M \equiv N$,

$$\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dM_s dM_t = \int_0^T \int_0^{t-} G(t-s) dM_s dM_t + \frac{G(0)}{2} \sum_{t \in [0, T]} \Delta M_t \Delta M_t. \quad (8.12)$$

Using (8.11) with $M_s = \tilde{\mathbb{E}}[X_s^i]$ and $N_t = X_t^i$ and then applying Theorem 3.2 yields

$$\begin{aligned} &\mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] dX_t^i \right] \\ &= \mathbb{E} \left[2 \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] dX_t^i + G(0) \sum_{t \in [0, T]} \Delta \tilde{\mathbb{E}}[X_t^i] \Delta X_t^i \right. \\ &\quad \left. + \int_0^T \int_0^{t-} G(t-s) dX_s^i d\tilde{\mathbb{E}}[X_t^i] - \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] dX_t^i \right] \\ &= \mathbb{E} \left[2 \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d(X_t^i)^{\mathbb{P}} + G(0) \sum_{t \in [0, T]} \Delta \tilde{\mathbb{E}}[X_t^i] \Delta (X_t^i)^{\mathbb{P}} \right. \\ &\quad \left. + \int_0^T \mathbb{P} \left(\int_0^{t-} G(t-s) dX_s^i \right) d\tilde{\mathbb{E}}[X_t^i] - \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d(X_t^i)^{\mathbb{P}} \right] \\ &= \mathbb{E} \left[2 \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] + G(0) \sum_{t \in [0, T]} \left(\Delta \tilde{\mathbb{E}}[X_t^i] \right)^2 \right. \\ &\quad \left. + \int_0^T \mathbb{P} \left(\int_0^{t-} G(t-s) dX_s^i \right) d\tilde{\mathbb{E}}[X_t^i] - \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right]. \end{aligned}$$

As seen in the proof of Proposition 4.2,

$$\mathbb{P} \left(\int_0^{t-} G(t-s) dX_s^i \right) = \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i].$$

Substituting this, simplifying, and applying (8.12) with $M_s = \tilde{\mathbb{E}}[X_s^i]$,

$$\begin{aligned}
& \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] dX_t^i \right] \\
&= \mathbb{E} \left[2 \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] + G(0) \sum_{t \in [0, T]} \left(\Delta \tilde{\mathbb{E}}[X_t^i] \right)^2 \right. \\
&\quad \left. + \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] - \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right] \\
&= \mathbb{E} \left[2 \int_0^T \int_0^{t-} G(t-s) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] + G(0) \sum_{t \in [0, T]} \left(\Delta \tilde{\mathbb{E}}[X_t^i] \right)^2 \right] \\
&= \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right]. \tag{8.13}
\end{aligned}$$

Substituting (8.13) into (8.10) and simplifying completes the proof. \square

Next, we prove the main result.

Proof of Theorem 4.5. Using Proposition 4.2 and (2.7) and the definition of the dual predictable projection in Theorem 3.2,

$$\begin{aligned}
J(X^i; \mathbf{X}^{-i}) &= J(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}]) \\
&= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j] d(X_t^i)^{\mathbb{P}} \right. \\
&\quad \left. + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta \tilde{\mathbb{E}}[X_t^j] \Delta(X_t^i)^{\mathbb{P}} + C(X^i) \right] \\
&= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j] d\tilde{\mathbb{E}}[X_t^i] \right. \\
&\quad \left. + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta \tilde{\mathbb{E}}[X_t^j] \Delta \tilde{\mathbb{E}}[X_t^i] + C(X^i) \right], \tag{8.14}
\end{aligned}$$

where the last equality used, similarly as in the proof of Proposition 4.2, that ${}^{\mathbb{P}}X^i = (X^i)^{\mathbb{P}} = \tilde{\mathbb{E}}[X^i]$ by Proposition 3.4 and the rule ${}^{\mathbb{P}}(\Delta A) = \Delta(A^{\mathbb{P}})$.

By our assumption, X^i cannot be equal to $\tilde{\mathbb{E}}[X^i]$ up to evanescence. As $x^i = X_{0-}^i = \tilde{\mathbb{E}}[X_{0-}^i]$, this implies that the random measure $d(X^i - \tilde{\mathbb{E}}[X^i])$ is non-zero with positive probability. Recalling that G is strict positive definite and invoking Lemma 8.5, it follows that

$$\begin{aligned}
0 &< \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) d(X_s - \tilde{\mathbb{E}}[X_s]) d(X_t - \tilde{\mathbb{E}}[X_t]) \right] \\
&= \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s dX_t - \frac{1}{2} \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s] d\tilde{\mathbb{E}}[X_t] \right]
\end{aligned}$$

which after rearranging states that

$$\mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i \right] > \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right].$$

Combining this with (8.14), we obtain

$$\begin{aligned} J(X^i; \mathbf{X}^{-i}) &> \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j] d\tilde{\mathbb{E}}[X_t^i] \right. \\ &\quad \left. + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta \tilde{\mathbb{E}}[X_t^j] \Delta \tilde{\mathbb{E}}[X_t^i] + C(X^i) \right]. \end{aligned}$$

By Assumption 4.3 we have $\mathbb{E}[C(X^i)] \geq \mathbb{E}[C(\tilde{\mathbb{E}}[X^i])]$. Thus we may replace $C(X^i)$ by $C(\tilde{\mathbb{E}}[X^i])$ in the above inequality, and then the right hand side is precisely $J(\tilde{\mathbb{E}}[X^i]; \mathbb{E}[\mathbf{X}^{-i}])$. Applying Proposition 4.2 once more yields the claim:

$$J(X^i; \mathbf{X}^{-i}) > J(\tilde{\mathbb{E}}[X^i]; \tilde{\mathbb{E}}[\mathbf{X}^{-i}]) = J(\tilde{\mathbb{E}}[X^i]; \mathbf{X}^{-i}). \quad \square$$

8.5 Proofs for Section 5 (Uniqueness)

The first step towards uniqueness of Nash equilibria is to ensure that the traders' objectives are convex in their controls. (In fact, the proof of our main result below only uses the assertion of convexity, not the strict convexity provided in the lemma.)

Lemma 8.6. *For any admissible \mathbf{X}^{-i} , the objective $J(\cdot; \mathbf{X}^{-i})$ is strictly¹² convex in its first argument.*

The proof is analogous to the proof in [21, Lemma 4.7] for the exponential kernel and a quadratic cost C . We include it for completeness and to clarify that it relies only on the convexity of the cost C and the strict positive definiteness of the kernel G .

Proof. Let Y^0 and Y^1 be distinct admissible strategies for trader i with $\mathbb{E}[C(Y^0)] < \infty$ and $\mathbb{E}[C(Y^1)] < \infty$. Fix any admissible strategy profile \mathbf{X}^{-i} of the other traders. For any $\alpha \in (0, 1)$ define

$$Y^\alpha := \alpha Y^1 + (1 - \alpha) Y^0$$

so that by the convexity of C ,

$$\begin{aligned} J(Y^\alpha; \mathbf{X}^{-i}) &\leq \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dY_s^\alpha dY_t^\alpha + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^j dY_t^\alpha \right. \\ &\quad \left. + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta X_t^j \Delta Y_t^\alpha + \alpha C(Y^1) + (1 - \alpha) C(Y^0) \right]. \end{aligned}$$

¹²An extended real-valued convex function F is called strictly convex if it is strictly convex on its domain $\text{dom}(F) = \{h : F(h) < \infty\}$.

Observe that the middle two terms are linear in α , so to show strict convexity of $J(\cdot; \mathbf{X}^{-i})$ it suffices to show

$$H(Y^\alpha) := \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^\alpha dY_t^\alpha \right] < \alpha \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^1 dY_t^1 \right] \\ + (1-\alpha) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^0 dY_t^0 \right]. \quad (8.15)$$

Expanding the left hand side of (8.15),

$$H(Y^\alpha) = \alpha^2 \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_t^1 dY_t^1 \right] + (1-\alpha)^2 \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^0 dY_t^0 \right] \\ + 2\alpha(1-\alpha) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^1 dY_t^0 \right]. \quad (8.16)$$

At the same time, by the strict positive definiteness of G ,

$$0 < \alpha(1-\alpha) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) d(Y_s^1 - Y_s^0) d(Y_t^1 - Y_t^0) \right] \\ = \alpha(1-\alpha) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^1 dY_t^1 \right] + \alpha(1-\alpha) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^0 dY_t^0 \right] \\ - 2\alpha(1-\alpha) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^1 dY_t^0 \right]. \quad (8.17)$$

Combining (8.16) and (8.17) we obtain

$$H(Y^\alpha) < (\alpha^2 + \alpha(1-\alpha)) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_t^1 dY_t^1 \right] \\ + ((1-\alpha)^2 + \alpha(1-\alpha)) \mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) dY_s^0 dY_t^0 \right].$$

Simplifying the coefficients recovers (8.15) and completes the proof. \square

Proof of Theorem 5.1. Let \mathbf{X}^0 and \mathbf{X}^1 be distinct equilibria. Define

$$\mathbf{X}^\alpha = \alpha \mathbf{X}^1 + (1-\alpha) \mathbf{X}^0, \quad \alpha \in (0, 1),$$

and

$$V(\alpha) = \sum_{i=1}^N (J(X^{\alpha,i}; \mathbf{X}^{0,-i}) + J(X^{1-\alpha,i}; \mathbf{X}^{1,-i})).$$

It follows from Lemma 8.6 is that V is strictly convex. Moreover, the Nash equilibrium property implies that

$$V(\alpha) \geq \sum_{i=1}^N (J(X^{0,i}; \mathbf{X}^{0,-i}) + J(X^{1,i}; \mathbf{X}^{1,-i})) = V(0). \quad (8.18)$$

Next, we will show that the right derivative at zero satisfies $\dot{V}(0+) < 0$. This will be a contradiction to (8.18) as it implies that 0 is not in the subdifferential of V at $\alpha = 0$ which, in turn, implies that $V(0)$ is not a minimum.

We treat separately the four terms in the expression (2.7) for J and their corresponding terms in V . Let

$$F_1^i(\alpha) = \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^{\alpha,i} dX_t^{\alpha,i} + \frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^{1-\alpha,i} dX_t^{1-\alpha,i} \right].$$

Substituting the explicit form of \mathbf{X}^α , $\mathbf{X}^{1-\alpha}$ and using the linearity of integration with respect to scalar multiplication, we see that $F_1^i(\alpha)$ is quadratic in α . Differentiating at $\alpha = 0$,

$$\dot{F}_1^i(0) = -\mathbb{E} \left[\int_0^T \int_0^T G(|t-s|) d(X_s^{1,i} - X_s^{0,i}) d(X_t^{1,i} - X_t^{0,i}) \right]$$

and thus

$$\sum_{i=1}^N \dot{F}_1^i(0) = -\mathbb{E} \left[\sum_{i=1}^N \int_0^T \int_0^T G(|t-s|) d(X_s^{1,i} - X_s^{0,i}) d(X_t^{1,i} - X_t^{0,i}) \right]. \quad (8.19)$$

Similarly, let

$$F_2^i(\alpha) = \mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^{0,j} dX_t^{\alpha,i} + \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^{1,j} dX_t^{1-\alpha,i} \right],$$

which is a linear function of α . Differentiating gives

$$\begin{aligned} \dot{F}_2^i(\alpha) &= \mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^{0,j} d(X_t^{1,i} - X_t^{0,i}) \right. \\ &\quad \left. - \int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} dX_s^{1,j} d(X_t^{1,i} - X_t^{0,i}) \right] \\ &= -\mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d(X_s^{1,j} - X_s^{0,j}) d(X_t^{1,i} - X_t^{0,i}) \right], \end{aligned}$$

which is constant in α . Summing over i yields

$$\begin{aligned}
\sum_{i=1}^N \dot{F}_2^i(0) &= -\mathbb{E} \left[\sum_{i=1}^N \sum_{j \neq i} \int_0^T \int_0^{t-} G(t-s) d(X_s^{1,j} - X_s^{0,j}) d(X_t^{1,i} - X_t^{0,i}) \right] \\
&= \mathbb{E} \left[-\frac{1}{2} \sum_{i=1}^N \sum_{j \neq i} \int_0^T \int_0^{t-} G(t-s) d(X_s^{1,j} - X_s^{0,j}) d(X_t^{1,i} - X_t^{0,i}) \right. \\
&\quad \left. - \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i} \int_0^T \int_{s+}^T G(t-s) d(X_t^{1,i} - X_t^{0,i}) d(X_s^{1,j} - X_s^{0,j}) \right] \\
&= \mathbb{E} \left[-\frac{1}{2} \sum_{i=1}^N \sum_{j \neq i} \int_0^T \int_0^T G(|t-s|) d(X_s^{1,j} - X_s^{0,j}) d(X_t^{1,i} - X_t^{0,i}) \right. \\
&\quad \left. + \frac{G(0)}{2} \sum_{i=1}^N \sum_{j \neq i} \sum_{t \in [0, T]} \Delta(X_t^{1,j} - X_t^{0,j}) \Delta(X_t^{1,i} - X_t^{0,i}) \right].
\end{aligned}$$

As G is positive definite (in fact, here we only use semi-definiteness),

$$-\sum_i \sum_{j \neq i} \int_0^T \int_0^T G(|t-s|) dM_s^j dM_t^i \leq \sum_i \int_0^T \int_0^T G(|t-s|) dM_s^i dM_t^i$$

for arbitrary Lebesgue–Stieltjes integrators M^i . Applying this,

$$\begin{aligned}
\sum_{i=1}^N \dot{F}_2^i(0) &\leq \mathbb{E} \left[\frac{1}{2} \sum_{i=1}^N \int_0^T \int_0^T G(|t-s|) d(X_s^{1,i} - X_s^{0,i}) d(X_t^{1,i} - X_t^{0,i}) \right. \\
&\quad \left. + \frac{G(0)}{2} \sum_{i=1}^N \sum_{j \neq i} \sum_{t \in [0, T]} \Delta(X_t^{1,j} - X_t^{0,j}) \Delta(X_t^{1,i} - X_t^{0,i}) \right]. \tag{8.20}
\end{aligned}$$

Turning to the jump terms, let

$$F_3^i(\alpha) = \mathbb{E} \left[\frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta X_t^{\alpha, i} \Delta X_t^{0, j} + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta X_t^{1-\alpha, i} \Delta X_t^{1, j} \right].$$

This is once again linear in α and differentiating gives

$$\dot{F}_3^i(\alpha) = \mathbb{E} \left[-\frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta(X_t^{0, i} - X_t^{1, i}) \Delta X_t^{0, j} + \frac{G(0)}{2} \sum_{j \neq i} \sum_{t \in [0, T]} \Delta(X_t^{0, i} - X_t^{1, i}) \Delta X_t^{1, j} \right].$$

Collecting terms and summing over i ,

$$\sum_{i=1}^N \dot{F}_3^i(0) = -\mathbb{E} \left[\frac{G(0)}{2} \sum_{i=1}^N \sum_{j \neq i} \sum_{t \in [0, T]} \Delta(X_t^{1, i} - X_t^{0, i}) \Delta(X_t^{1, j} - X_t^{0, j}) \right]. \tag{8.21}$$

Lastly, we treat the terms arising from the cost C . Let

$$F_4^i(\alpha) = \mathbb{E} [C(X^{\alpha,i}) + C(X^{1-\alpha,i})].$$

By construction, $[0, 1] \ni \alpha \mapsto F_4^i(\alpha)$ is a finite convex function and so its right derivative exists at 0 (we allow for the possibility that $\dot{F}_4^i(0+) = -\infty$). Moreover, $F_4^i(\alpha)$ is symmetric about $\alpha = 1/2$ which together with convexity implies that $\alpha = 1/2$ is a minimum. Hence, $0 \in \partial F_4^i(1/2)$ where $\partial F_4^i(\alpha)$ denotes the subgradient set at α . By convexity of F_4^i and the definition of the subgradient, $\dot{F}_4^i(0+) \leq y$ for any $y \in \partial F_4^i(1/2)$. Taking $y = 0$ and summing over i we conclude

$$\sum_{i=1}^N \dot{F}_4^i(0+) \leq 0. \tag{8.22}$$

Aggregating the above expressions we recover

$$V(\alpha) = \sum_{i=1}^N (F_1^i(\alpha) + F_2^i(\alpha) + F_3^i(\alpha) + F_4^i(\alpha))$$

and obtain its right derivative at 0,

$$\begin{aligned} \dot{V}(0+) &= \sum_{i=1}^N \left(\dot{F}_1^i(0) + \dot{F}_2^i(0) + \dot{F}_3^i(0) + \dot{F}_4^i(0+) \right) \\ &\leq -\mathbb{E} \left[\frac{1}{2} \sum_{i=1}^N \int_0^T \int_0^T G(|t-s|) d(X_s^{1,i} - X_s^{0,i}) d(X_t^{1,i} - X_t^{0,i}) \right] < 0. \end{aligned}$$

Here the first inequality follows from Eqs. (8.19) to (8.22) whereas the last inequality holds as G is strictly positive definite and $\mathbf{X}^0 \neq \mathbf{X}^1$. As mentioned above, $\dot{V}(0+) < 0$ is a contradiction and completes the proof. \square

8.6 Proofs for Section 6 (Singular Kernels)

In this subsection, we extend our results to the case of a singular kernel, where ‘singular kernel’ will refer to a function G satisfying Assumption 6.1. Apart from technical verifications, the extension is based on the fact that a singular kernel admits a monotone approximation by kernels satisfying Assumption 2.3; cf. Remark 6.5.

8.6.1 Proof of Lemma 6.4

Proof of Lemma 6.4. We need to show that

$$\int_0^T \int_0^T G(|t-s|) dX_s dX_t > 0$$

whenever $X : [0, T] \rightarrow \mathbb{R}$ is a càdlàg function of finite variation that is not constant and satisfies $\int_0^T \int_0^T G(|t-s|) d|X|_s d|X|_t < \infty$. By [13, Proposition 4.5], the double integral admits a representation of the form

$$\int_0^T \int_0^T G(|t-s|) dX_s dX_t = \int |\widehat{X}(z)|^2 \mu(dz). \quad (8.23)$$

As X has finite variation, the Fourier–Stieltjes transform $\widehat{X}(z) = \int_0^T e^{itz} dX_t$ is a continuous function that does not vanish identically as soon as X is not constant. Moreover, [13, Lemma 4.2] gives an expression for the positive Radon measure μ from which we see that μ has full support. As a consequence, the right hand side of (8.23) is strictly positive. \square

8.6.2 Objective Function (Proposition 6.6)

We first record a simple lemma for later of reference. Recall Definition 6.2 of admissibility.

Lemma 8.7. *For a profile $\mathbf{X} = (X^1, \dots, X^N)$ of admissible strategies and $i, j \in \{1, \dots, N\}$,*

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|X^j|_s d|X^i|_t \right] \\ \leq \frac{1}{2} \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) (d|X^i|_s d|X^i|_t + d|X^j|_s d|X^j|_t) \right] < \infty. \end{aligned}$$

Proof. Using the positive definiteness of G ,

$$\begin{aligned} 0 &\leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d(|X^i|_s - |X^j|_s) d(|X^i|_t - |X^j|_t) \right] \\ &= \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) (d|X^i|_s d|X^i|_t + d|X^j|_s d|X^j|_t - 2d|X^i|_s d|X^j|_t) \right]. \end{aligned}$$

Rearranging and applying the admissibility of X^i and X^j yields the claim. \square

Proof of Proposition 6.6. Recalling the absence of jumps in the present setting (Remark 6.3), the proof follows the same lines as the one of Proposition 2.6; we merely indicate where additional justifications are needed. First, to justify the application of Fubini’s theorem below (8.1), we observe that

$$\left| \int_0^T \int_0^{t-} G(t-s) \sum_{j=1}^N dX_s^j dX_t^i \right| \leq \int_0^\infty \int_0^\infty G(|t-s|) \sum_{j=1}^N d|X^j|_s d|X^i|_t$$

where the right hand side is integrable and therefore almost surely finite by Lemma 8.7. Second, to argue $|J(X^i; \mathbf{X}^{-i})| < \infty$ at the end of the proof, note that (6.2) gives

$$\begin{aligned} |J(X^i; \mathbf{X}^{-i})| \\ \leq \mathbb{E} \left[\frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) d|X^i|_s d|X^i|_t + \int_0^\infty \int_0^\infty G(|t-s|) \sum_{j \neq i} d|X^j|_s d|X^i|_t + C(X^i) \right] \end{aligned}$$

which is finite by admissibility, Lemma 8.7 and our assumption that $\mathbb{E}[C(X^i)] < \infty$. \square

8.6.3 Extension of Lemma 4.1

Let X be an admissible strategy for the singular kernel G (Definition 6.2). To extend Lemma 4.1 to the case where G is singular, it suffices to show that $\tilde{\mathbb{E}}[X]$ satisfies the integrability condition (6.1).

Clearly the processes $t \mapsto TV(X; [0, t])$, $t \mapsto TV(\tilde{\mathbb{E}}[X]; [0, t])$ and $t \mapsto \tilde{\mathbb{E}}[TV(X; [0, t])]$ are càdlàg, increasing, and essentially bounded. (As X is continuous in the present setting, these processes are even continuous.) As above, we write $d|X|_t$ to denote integration with respect to the total variation of X and $d|\tilde{\mathbb{E}}[X]|_t$ to denote integration with respect to the total variation of $\tilde{\mathbb{E}}[X]$. Similarly, we write $d\tilde{\mathbb{E}}[|X|_t]$ to denote integration with respect to the measure induced by $t \mapsto \tilde{\mathbb{E}}[TV(X; [0, t])]$.

Lemma 8.8. *If G is a kernel satisfying Assumption 2.3,*

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t \right] &\leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d\tilde{\mathbb{E}}[|X|_s] d\tilde{\mathbb{E}}[|X|_t] \right] \\ &\leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|X|_s d|X|_t \right]. \end{aligned}$$

Proof. By Lemma 8.4 the non-negative measures corresponding to $TV(\tilde{\mathbb{E}}[X]; [0, t])$ and $\tilde{\mathbb{E}}[TV(X; [0, t])]$ are ordered.¹³ Therefore, as G is non-negative,

$$\int_0^\infty \int_0^\infty G(|t-s|) d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t \leq \int_0^\infty \int_0^\infty G(|t-s|) d\tilde{\mathbb{E}}[|X|_s] d\tilde{\mathbb{E}}[|X|_t],$$

from which the first claim follows by taking expectations. For the second inequality we use the positive definiteness of the kernel and apply Lemma 8.5 to the admissible process $t \mapsto TV(X; [0, t])$,

$$\begin{aligned} 0 &\leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d(|X|_s - \tilde{\mathbb{E}}[|X|_s]) d(|X|_t - \tilde{\mathbb{E}}[|X|_t]) \right] \\ &= \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|X|_s d|X|_t - \int_0^\infty \int_0^\infty G(|t-s|) d\tilde{\mathbb{E}}[|X|_s] d\tilde{\mathbb{E}}[|X|_t] \right]. \end{aligned}$$

Rearranging completes the proof. □

We are now in a position to prove that (6.1) holds for $\tilde{\mathbb{E}}[X^i]$.

Lemma 8.9. *If X is admissible for the singular kernel G , then*

$$\mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t \right] \leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|) d|X|_s d|X|_t \right] < \infty$$

and therefore $\tilde{\mathbb{E}}[X^i]$ is admissible; i.e., Lemma 4.1 holds for the singular kernel G .

¹³In the pointwise sense: two measures μ, ν are called ordered if $\mu(A) \leq \nu(A)$ for every Borel set A .

Proof. As seen in Remark 6.5, the singular kernel G can be approximated by a sequence of kernels $G^n \uparrow G$ satisfying Assumption 2.3. By the monotonicity of the sequence and the non-negativity of the integrating measures,

$$0 \leq \int_0^\infty \int_0^\infty G^n(|t-s|)d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t \leq \int_0^\infty \int_0^\infty G^{n+1}(|t-s|)d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t, \quad \forall n \geq 1.$$

Given that X is admissible for G , by applying Lemma 8.8 for each G^n we have

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty \int_0^\infty G^n(|t-s|)d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t \right] &\leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G^n(|t-s|)d|X|_s d|X|_t \right] \\ &\leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|)d|X|_s d|X|_t \right] < \infty. \end{aligned}$$

Using that $G^n \uparrow G$ for $d|\tilde{\mathbb{E}}[X]| \otimes d|\tilde{\mathbb{E}}[X]|$ almost every (s, t) , almost surely, we have by monotone convergence that

$$\mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|)d|\tilde{\mathbb{E}}[X]|_s d|\tilde{\mathbb{E}}[X]|_t \right] \leq \mathbb{E} \left[\int_0^\infty \int_0^\infty G(|t-s|)d|X|_s d|X|_t \right] < \infty. \quad \square$$

8.6.4 Extension of Proposition 4.2

The following stability property of the objective with respect to monotone approximation of G will be used to extend Proposition 4.2 to the singular case. When there are several kernels under consideration, we write J_G for the objective function with kernel G .

Proposition 8.10. *Let \mathbf{X} be an admissible strategy profile for the singular kernel G and let $(G^n)_{n \geq 1}$ be a sequence of kernels satisfying Assumption 2.3 with $G^n \uparrow G$. Then*

$$\lim_{n \rightarrow \infty} J_{G^n}(X^i; \mathbf{X}^{-i}) = J_G(X^i; \mathbf{X}^{-i}).$$

Proof. As \mathbf{X} is admissible for G , it has no jumps and the objective is given by (6.2). Consider $J_{G^n}(X^i; \mathbf{X}^{-i})$ as given by (6.2) with G^n , we want to take the limit and obtain the same expression with G instead of G^n . Indeed, by Lemma 8.7,

$$\frac{1}{2} \int_0^T \int_0^T G(|t-s|)d|X^i|_s d|X^i|_t + \int_0^T \int_0^T G(|t-s|) \sum_{j \neq i} d|X^j|_s d|X^i|_t$$

is integrable. As $0 \leq G^n \leq G$ and $G^n \rightarrow G$ pointwise, dominated convergence yields the claim. \square

Using this stability result, we readily obtain the extension of Proposition 4.2.

Proposition 8.11. *Proposition 4.2 holds for the singular kernel G .*

Proof. Given an admissible strategy profile \mathbf{X} for the singular kernel G , we need to show that $J_G(X^i; \mathbf{X}^{-i}) = J_G(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}])$. Define the approximating sequence $G^n \uparrow G$ as in Remark 6.5. We apply Proposition 8.10 to the strategies X^i, \mathbf{X}^{-i} and $\tilde{\mathbb{E}}[\mathbf{X}^{-i}]$ which are admissible for G by Lemma 8.9: using Proposition 4.2 for each G^n , we obtain

$$J_G(X^i; \mathbf{X}^{-i}) = \lim_{n \rightarrow \infty} J_{G^n}(X^i; \mathbf{X}^{-i}) = \lim_{n \rightarrow \infty} J_{G^n}(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}]) = J_G(X^i; \tilde{\mathbb{E}}[\mathbf{X}^{-i}]). \quad \square$$

8.6.5 Extension of Theorem 4.5

The extension of Theorem 4.5 does not follow from the stability result alone (which would only give a non-strict inequality). Instead, we appeal once again to strict positive definiteness.

Proposition 8.12. *Theorem 4.5 holds for the singular kernel G .*

Proof. In view of (6.2), it suffices to show two (in)equalities,

$$\mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j] dX_t^i \right] = \mathbb{E} \left[\int_0^T \int_0^{t-} G(t-s) \sum_{j \neq i} d\tilde{\mathbb{E}}[X_s^j] d\tilde{\mathbb{E}}[X_t^i] \right] \quad (8.24)$$

and

$$\mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) dX_s^i dX_t^i \right] > \mathbb{E} \left[\frac{1}{2} \int_0^T \int_0^T G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right]. \quad (8.25)$$

Define again the approximating sequence $G^n \uparrow G$ as in Remark 6.5. Then (8.24) holds with G^n instead of G , and since X^i, X^j satisfy the admissibility condition (6.1), we can repeat the same argument as in Proposition 8.10 to apply dominated convergence and deduce (8.24) for G . To prove (8.25), we first argue that the equality

$$\begin{aligned} \mathbb{E} \left[\frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) d(X_s^i - \tilde{\mathbb{E}}[X_s^i]) d(X_t^i - \tilde{\mathbb{E}}[X_t^i]) \right] \\ = \mathbb{E} \left[\frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) dX_s^i dX_t^i - \frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) d\tilde{\mathbb{E}}[X_s^i] d\tilde{\mathbb{E}}[X_t^i] \right] \end{aligned} \quad (8.26)$$

from Lemma 8.5 extends to singular G . Indeed, (8.26) holds for each G^n by Lemma 8.5, and given that X^i and $\tilde{\mathbb{E}}[X^i]$ satisfy (6.1), dominated convergence then yields (8.26) for G . As G is strictly positive definite (Lemma 6.4), our assumption that $X^i - \mathbb{E}[X^i]$ is not identically zero yields

$$0 < \mathbb{E} \left[\frac{1}{2} \int_0^\infty \int_0^\infty G(|t-s|) d(X_s^i - \tilde{\mathbb{E}}[X_s^i]) d(X_t^i - \tilde{\mathbb{E}}[X_t^i]) \right].$$

Applying (8.26) and rearranging gives the desired inequality (8.25). \square

8.6.6 Proof of Theorem 6.8

In the preceding subsections, we have shown the extensions of Lemma 4.1, Proposition 4.2 and Theorem 4.5. The extension of the uniqueness result (Theorem 5.1) is obtained by following the steps in the proof of Theorem 5.1 and omitting the jump terms. To justify the applications of Fubini's theorem, we use Lemma 8.7 as in the proof of Proposition 6.6.

References

- [1] E. Abi Jaber, E. Neuman, and M. Voß. Equilibrium in functional stochastic games with mean-field interaction. *Preprint arXiv:2306.05433v1*, 2024.

- [2] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM J. Financial Math.*, 3(1):511–533, 2012.
- [3] S. Campbell and M. Nutz. Optimal execution among N traders with transient price impact. *Preprint arXiv:2501.09638v1*, 2025.
- [4] B. I. Carlin, M. S. Lobo, and S. Viswanathan. Episodic liquidity crises: Cooperative and predatory trading. *J. Finance*, 62(5):2235–2274, 2007.
- [5] Á. Cartea, S. Jaimungal, and J. Penalva. *Algorithmic and High-Frequency Trading*. Cambridge University Press, 2015.
- [6] P. Casgrain and S. Jaimungal. Mean field games with partial information for algorithmic trading. *Preprint arXiv:1803.04094v2*, 2019.
- [7] P. Casgrain and S. Jaimungal. Mean-field games with differing beliefs for algorithmic trading. *Math. Finance*, 30(3):995–1034, 2020.
- [8] C. Dellacherie and P. A. Meyer. *Probabilities and Potential B*. North Holland, Amsterdam, 1982.
- [9] G. Fu, U. Horst, and X. Xia. Portfolio liquidation games with self-exciting order flow. *Math. Finance*, 32(4):1020–1065, 2022.
- [10] N. Garleanu and L. H. Pedersen. Dynamic portfolio choice with frictions. *J. Econ. Theory*, 165:487–516, 2016.
- [11] J. Gatheral. No-dynamic-arbitrage and market impact. *Quant. Finance*, 10(7):749–759, 2010.
- [12] J. Gatheral and A. Schied. Dynamical models of market impact and algorithms for order execution. In *Handbook on Systemic Risk*, pages 579–602. Cambridge University Press, 2013.
- [13] J. Gatheral, A. Schied, and A. Slynko. Transient linear price impact and Fredholm integral equations. *Math. Finance*, 22(3):445–474, 2012.
- [14] P. Graewe and U. Horst. Optimal trade execution with instantaneous price impact and stochastic resilience. *SIAM J. Control Optim.*, 55(6):3707–3725, 2017.
- [15] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, Berlin, 2nd edition, 2003.
- [16] C. C. Moallemi, B. Park, and B. Van Roy. Strategic execution in the presence of an uninformed arbitrageur. *J. Financ. Markets*, 15(4):361–391, 2012.
- [17] E. Neuman and M. Voß. Optimal signal-adaptive trading with temporary and transient price impact. *SIAM J. Financial Math.*, 13(2):551–575, 2022.
- [18] E. Neuman and M. Voß. Trading with the crowd. *Math. Finance*, 33(3):548–617, 2023.
- [19] A. A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *J. Financial Mark.*, 16(1):1–32, 2013.
- [20] L. H. Pedersen and M. K. Brunnermeier. Predatory trading. *J. Finance*, 60(4):1825–1863, 2005.
- [21] A. Schied, E. Strehle, and T. Zhang. High-frequency limit of Nash equilibria in a market impact game with transient price impact. *SIAM J. Financial Math.*, 8(1):589–634, 2017.
- [22] A. Schied and T. Zhang. A market impact game under transient price impact. *Math. Oper. Res.*, 44(1):102–121, 2019.
- [23] T. Schöneborn. *Trade execution in illiquid markets*. PhD thesis, TU Berlin, 2008.
- [24] T. Schöneborn and A. Schied. Liquidation in the face of adversity: Stealth vs. sunshine trading. *Preprint SSRN:1007014*, 2009.
- [25] E. Strehle. Optimal execution in a multiplayer model of transient price impact. *Market Microstructure and Liquidity*, 03(03n04):1850007, 2017.
- [26] K. Webster. *Handbook of Price Impact Modeling*. CRC Press, Boca Raton, FL, 2023.