

# A characterisation of cross-impact kernels \*

Mathieu Rosenbaum<sup>†</sup> Mehdi Tomas<sup>‡</sup>

July 20, 2021

## Abstract

Trading a financial asset pushes its price as well as the prices of other assets, a phenomenon known as cross-impact. We consider a general class of kernel-based cross-impact models and investigate suitable parameterisations for trading purposes. We focus on kernels that guarantee that prices are martingales and anticipate future order flow (martingale-admissible kernels) and those that ensure there is no possible price manipulation (no-statistical-arbitrage-admissible kernels). We determine the overlap between these two classes and provide formulas for calibration of cross-impact kernels on data. We illustrate our results using SP500 futures data.

**Keywords:** Cross impact, market impact, multidimensional processes, market microstructure, market efficiency, statistical arbitrage

**AMS 2000 subject classifications:** 60G44, 60G55, 62M10.

## Introduction

How do trades move prices of financial securities? It is well-known among practitioners and academics that buying a financial asset tends to push its price up while selling it tends to push its price down. This observation is one aspect of *market impact*, which describes how trades on one asset translate into its price. The many studies on market impact [2, 6, 25] have deepened our understanding of how markets digest trades into prices. In turn, this has helped us understand key properties of dynamics of asset prices. For instance, market impact explains why price volatilities are well-modeled by rough fractional Brownian motions [15].

Yet classical market impact does not tell us how our trades influence prices of other assets. Thus, it ignores a potentially important aspect of price formation. As many assets are simultaneously traded in financial markets, this element is required to generate complete market dynamics and, ultimately, answer the question of how markets digest liquidity.

---

\*Mathieu Rosenbaum and Mehdi Tomas gratefully acknowledge financial support of the ERC 679836 StaQamof and the Chair Analytics and Models for financial regulation, Deep finance and statistics, Machine learning and systematic methods. Mehdi Tomas gratefully acknowledges the *Econophysics & Complex Systems* Research Chair under the aegis of the Fondation du Risque, a joint initiative by the *Fondation de l'École polytechnique*, *l'École polytechnique* and Capital Fund Management. The authors thank Michael Benzaquen, Iacopo Mastromatteo and Michele Vodret for helpful discussions and comments.

<sup>†</sup>CMAP, École Polytechnique, mathieu.rosenbaum@polytechnique.edu

<sup>‡</sup>CMAP & LadHyx, École Polytechnique, mehdi.tomas@polytechnique.edu

To solve this issue, a recent strand of works [1, 5, 20, 22, 23, 24, 28, 30] has studied *cross-impact*, which describes how transactions on a universe of instruments drive their prices. This paper contributes to the literature by characterising cross-impact models which lead to well-behaved market dynamics. We show how these models can be calibrated from empirical data and provide an example using SP500 futures.

We consider a stylised market made of  $d$  financial securities, continuously quoted and traded by market participants. Trading activity on financial markets is highly endogenous: statistically, trades trigger other trades [6, 11]. To capture this effect, we model trade dynamics in our market by Hawkes processes [13] and introduce the  $d$ -dimensional counting processes  $N_t^a/N_t^b$  denoting the number of buy/sell orders of all market participants over  $[0, t]$ . We classically assume that trades induce permanent impact which is linear in the traded volume, for reasons related to no-arbitrage [10, 20]. A large class of models which satisfy this property and has been well-studied when  $d = 1$  is the class of propagator models [6]. Thus we restrict ourselves to this class and assume that  $d$ -dimensional price process  $P$  evolves as

$$P_t = P_0 + \int_0^t K(t-s)(dN_s^a - dN_s^b), \quad (1)$$

where  $K: t \in \mathbb{R}_+ \mapsto K(t) \in \mathcal{M}_d(\mathbb{R})$  is a *cross-impact kernel*. The cross-impact kernel encodes all information about cross-impact in our market but, contrary to prices and trades, it is not directly observable. When it exists, the limit of the cross-impact kernel  $\Lambda := \lim_{t \rightarrow \infty} K(t)$  is called the permanent cross-impact matrix since  $\Lambda_{ij}$  quantifies the permanent price impact of a trade on Asset  $j$  on the price of Asset  $i$ .

Within Equation (1), we examine two different classes of cross-impact kernels: those that anticipate upcoming order flow and yield martingale prices, which we dub *martingale-admissible*, and those that prevent statistical arbitrage, which we dub *no-statistical-arbitrage-admissible*, or *nsa-admissible* for short. Statistical arbitrage is meant in the sense of [10]: a statistical arbitrage is a trading strategy that starts and ends with no asset holdings and has negative expected costs.

Martingale-admissible and nsa-admissible kernels enforce different aspects of price efficiency. Martingale-admissible kernels ensure that prices are not predictable and that information flow is reflected in the current price, so that no trading strategy can make a profit by forecasting prices or order flows. On the other hand, nsa-admissible kernels prevent price manipulation by large agents who could push prices to make a profit.

The main contribution of this paper is the characterisation of the class of martingale-admissible and nsa-admissible kernels with respect to price and order flow statistics. In particular, at most one cross-impact kernel is both martingale-admissible and nsa-admissible. This characterisation can be used for calibration on real data and we provide an application on market data to illustrate our results.

We now comment on the links between our approach and the literature.

This market model using Hawkes processes in a propagator framework is a generalisation of the model from [14, 15], where only one asset is considered. In the single-asset case, this market dynamic is consistent with many empirical results concerning market impact. Therefore the multivariate generalisation of this framework will

stay consistent with these findings while providing insights into cross-impact.

The papers [5, 20] study a class of cross-impact kernels which give rise to martingale prices. This condition is used to obtain a calibration methodology. Though the calibration methodology based on maximum likelihood is straightforward, the resulting cross-impact kernels are noisy, prone to overfitting and they have no guarantees of no-arbitrage. We show that all martingale-admissible kernels, including those presented in [5, 20], have a certain form.

The class of nsa-admissible kernels has been described in [1]. However, the provided characterisation is quite theoretical and gives no insight into which cross-impact kernels to choose in practice. This paper extends some of the results of [1]. In particular, we show that nsa-admissible kernels have constrained values at zero and infinity which are related to price and order flow statistics. Unfortunately, there are many kernels which are nsa-admissible but which lead to ill-behaved market dynamics, as pointed out in [1]. We provide a methodology for obtaining a nsa-admissible kernel which is close to a martingale-admissible kernel. Doing so, we obtain a kernel which is still faithful to empirical data while preventing statistical arbitrage.

The resulting kernels can be used on market data to estimate cross-impact. Thus, the paper adds to the literature focusing on calibrating cross-impact kernels [5, 12, 19, 20, 29]. While some kernels are nsa-admissible such as the eigenliquidity cross-impact kernels [5], others are martingale-admissible [5, 20]. This paper provides kernels which can be easily calibrated in both classes.

Finally, the boundary values  $K(0)$  and  $\lim_{t \rightarrow \infty} K(t)$  of any cross-impact kernel  $K$  which is martingale-admissible and satisfies necessary conditions for nsa-admissibility have a microscopic foundation. Indeed, both can be interpreted as solutions to the multivariate version of Kyle's insider trading problem [9]. Thus, although cross-impact is purely a reaction to order flow imbalance in our model, the cross-impact kernel can also be interpreted through the lens of information revelation.

The paper is organized as follows. In Section 1, we describe our financial market. In Section 2, we characterise the classes of martingale-admissible and nsa-admissible kernels. Finally, we apply our results on market data in Section 3 before concluding in Section 4. Some proofs and additional results are relegated to an appendix.

## Notation

The set of  $d \times d$  real-valued square matrices is denoted by  $\mathcal{M}_d(\mathbb{R})$ , the set of orthogonal (also called rotation) matrices by  $\mathcal{O}_d(\mathbb{R})$ , the set of real symmetric matrices by  $\mathcal{S}_d(\mathbb{R})$  and the set of real symmetric positive matrices by  $\mathcal{S}_d^+(\mathbb{R})$ . Furthermore, given a matrix  $A$  in  $\mathcal{M}_d(\mathbb{R})$ ,  $A^\top$  denotes its transpose. Given  $A$  in  $\mathcal{S}_d^+(\mathbb{R})$ , we write  $A^{1/2}$  for a matrix such that  $A^{1/2}(A^{1/2})^\top = A$  and  $\sqrt{A}$  for the square root matrix, the unique positive semi-definite symmetric matrix such that  $(\sqrt{A})^2 = A$ . Finally, given a vector  $v \in \mathbb{R}^d$ , we write  $v = (v_1, \dots, v_d)$  and  $\text{diag}(v)$  for the diagonal matrix with entries the elements of  $v$ .

A matrix  $M \in \mathcal{M}_d(\mathbb{R})$  is called non-negative if for any  $x \in \mathbb{R}^d$ ,  $x^\top Mx \geq 0$ . It is called non-negative definite if  $z^* M z \geq 0$  for any  $z \in \mathbb{C}^d$ . A matrix  $M \in \mathcal{M}_d(\mathbb{C})$  is called strictly positive if  $x^\top Mx > 0$  for any nonzero  $x \in \mathbb{R}^d$  and

strictly positive definite if  $z^* M z > 0$  for any nonzero  $z \in \mathbb{C}^d$ . The conjugate transpose of a matrix  $M \in \mathcal{M}_d(\mathbb{C})$  is written  $M^*$ .

A function  $f: \mathbb{R} \rightarrow \mathbb{R}$  is called causal if  $f(t) = 0$  for all negative  $t$ . The same is said of a vector-valued or matrix-valued function if all its entries are causal. Given an integrable function  $f$ , we denote its Fourier transform  $\hat{f}: \mathbb{C} \rightarrow \mathbb{C}$  defined, for all  $\omega \in \mathbb{C}$  such that the integral converges, by

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt.$$

Similarly, we define the Fourier transform of a vector-valued or matrix-valued function with integrable entries by the vector-valued or matrix-valued function Fourier transform of all its entries. For a given real-valued measure  $\mu: \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ , we denote its Fourier transform by  $\hat{\mu}: \mathbb{C} \rightarrow \mathbb{C}$  defined, for all  $\omega \in \mathbb{C}$  such that the integral converges, by

$$\hat{\mu}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \mu(dt).$$

Similarly, we define the Fourier transform of a vector or matrix with measure entries by the Fourier transform of all its entries. All stochastic processes in the text are defined on a probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}}, \mathbb{P})$ . Given two semi-martingale processes  $X$  and  $Y$ , we denote by  $\langle X, X \rangle$  the predictable quadratic variation of  $X$  and  $\langle X, Y \rangle$  the predictable quadratic covariation of  $X$  and  $Y$ .

## 1 Market model

This section presents the stylised market model in force throughout the paper. Our setting extends [14, 15] to the multivariate case. For a lengthier discussion about these assumptions in the univariate case, we refer the reader to [14, 15].

We consider a market made of  $d$  different assets, quoted and traded continuously in time. We greatly simplify the market by abstracting away microstructural features and assume that agents can buy and sell Asset  $i$  at time  $t$  at the unique quoted price  $P_t^i$ . The  $d$ -dimensional price process is denoted by  $P := (P^1, \dots, P^d)$ . We assume that agents trade from time 0 onward. During this period, the cumulative traded volume by all agents at the ask (resp. bid) is denoted by  $V^a$  (resp.  $V^b$ ) and the net traded volume by  $V := V^a - V^b$ .

A key property of the order flow in financial markets is its persistence: the sign correlation of orders is slowly decaying in time [6]. Because of this effect, a particularly successful model for order flow dynamics is the Hawkes process which can capture self-excitation and cross-excitation across time and instruments [4, 13]. Thus we will assume that the order flow dynamics are given by a Hawkes process.

**Assumption 1** (Hawkes order flow). *The number of buy and sell market orders follows a Hawkes process  $(N^a, N^b)$ ,*

of intensity  $(\lambda^a, \lambda^b)$  and kernel  $\Phi = \begin{pmatrix} \Phi^{a/a} & \Phi^{a/b} \\ \Phi^{b/a} & \Phi^{b/b} \end{pmatrix}$  such that

$$\begin{aligned}\lambda_t^a &= \mu + \int_0^t \Phi^{a/a}(t-s) dN_s^a + \int_0^t \Phi^{a/b}(t-s) dN_s^b \\ \lambda_t^b &= \mu + \int_0^t \Phi^{b/a}(t-s) dN_s^a + \int_0^t \Phi^{b/b}(t-s) dN_s^b,\end{aligned}$$

where in the above

- the vector  $\mu \in \mathbb{R}_+^d$  is the exogenous intensity of buy and sell market orders;
- the entry-wise integrable matrix function  $\Phi^{a/a}: t \mapsto \Phi^{a/a}(t) \in \mathcal{M}_d(\mathbb{R}_+)$  (resp.  $\Phi^{b/b}: t \mapsto \Phi^{b/b}(t) \in \mathcal{M}_d(\mathbb{R}_+)$ ) encodes the endogenous contribution of past buy (resp. sell) market orders on the intensity of buy (resp. sell) market orders;
- the entry-wise integrable matrix function  $\Phi^{a/b}: t \mapsto \Phi^{a/b}(t) \in \mathcal{M}_d(\mathbb{R}_+)$  (resp.  $\Phi^{b/a}: t \mapsto \Phi^{b/a}(t) \in \mathcal{M}_d(\mathbb{R}_+)$ ) encodes the endogenous contribution of past sell (resp. buy) market orders on the intensity of buy (resp. sell) market orders.

We assume that the Hawkes parameters are such that  $\mathbb{E}[\lambda_t^a] = \mathbb{E}[\lambda_t^b]$  for any  $t$ . Each market order on Asset  $i$  is assumed to be of constant size  $v_i$  and the spectral radius of the  $L_1$  norm of the Hawkes kernel  $\Phi$  is assumed to be below one. The latter assumption allows us to define the stationary version of the Hawkes process (in fact stationary intensity).

This framework allows for rich multivariate dynamics since we can account for self-excitation and buy/sell interactions between different assets through  $\Phi_{ij}^{b/a}$  and  $\Phi_{ji}^{a/b}$ . We assumed that the market there are as many buy market orders than sell market orders on the small time scales of the market model.

While the order flow is persistent, the order flow at time  $t$  should not give information about the order flow at time  $t' \gg t$ . We formalize this in the next assumption.

**Assumption 2** (Finitely predictable orderflow). *For all  $t \geq 0$ ,  $\mathbb{E}[V_s^a - V_s^b \mid \mathcal{F}_t]$  converges in probability to some finite limit as  $s$  tends to infinity.*

Note that this assumption implies additional constraints on the Hawkes parameters.

In our market, prices are driven by the transactions of all agents. To exclude trivial arbitrages and keep the model simple, it is natural to assume that the permanent component of the cross-impact is a linear function of the order flow (see for example Corollary 3.7 of [20]) and to consider a propagator framework as explained in the introduction [6]. This leads to the following assumption.

**Assumption 3** (Price dynamics). *There exists some function  $K: t \in \mathbb{R}_+ \mapsto K(t) \in \mathcal{M}_d(\mathbb{R})$ , called a cross-impact kernel, such that the price process  $P$  satisfies, for all  $t \in \mathbb{R}_+$*

$$P_t = P_0 + \int_0^t K(t-s)(dN_s^a - dN_s^b), \tag{2}$$

and  $K(t) \xrightarrow[t \rightarrow \infty]{} \Lambda$ , where  $\Lambda$  is an invertible  $d \times d$  matrix called the permanent cross-impact matrix.

The diagonal functions  $K_{ii}$  of the cross-impact kernel relate past order flow on a security to its price. Off-diagonal elements  $K_{ij}$  relate past order flow of Asset  $j$  to the price of Asset  $i$ . The matrix  $\Lambda$  is called the permanent cross-impact matrix since  $\Lambda_{ij}$  quantifies how much the price of asset  $i$  is moved by the net order flow on asset  $j$  after a long period.

Finally, we make a technical assumption about the continuity of the Hawkes kernel at the origin and its decay at infinity.

**Assumption 4.** *The Hawkes order flow kernel  $\Phi$  is continuously differentiable at zero and square-integrable.*

This assumption is not really constraining since Hawkes kernels for order flows are found to be square-integrable when calibrated on financial data [11].

Previous hypotheses have not touched on the efficiency of prices in our market. Without imposing additional assumptions, prices may be highly predictable or agents could manipulate them through trading to generate profits. These two concepts of price efficiency are not always compatible, even in this stylised model. As such, we need to distinguish between cross-impact kernels which give martingale prices and those that prevent statistical arbitrage.

The class of cross-impact kernels which give rise to martingale prices also includes trivial examples, such as  $K = 0$ . To exclude these, we introduce *martingale-admissible* kernels, which anticipate the impact contribution of the order flow and lead to martingale prices. Such kernels generate non-trivial price dynamics since they incorporate the impact contribution of trades in prices. In the univariate case  $d = 1$ , Theorem 2.1 of [14] shows that when prices are martingales and trades impact prices, we have

$$P_t - P_0 = \kappa \lim_{s \rightarrow \infty} \mathbb{E}[V_s^a - V_s^b | \mathcal{F}_t],$$

where  $\kappa > 0$  is the permanent market impact contribution. This motivates the following definition for *martingale-admissible* kernels.

**Definition 1.1** (Martingale-admissible kernels). *A cross-impact kernel  $K$  is said to be martingale-admissible if*

$$P_t - P_0 = \int_0^t K(t-s)(dN_s^a - dN_s^b) = \Lambda \lim_{s \rightarrow \infty} \mathbb{E}[V_s^a - V_s^b | \mathcal{F}_t]. \quad (3)$$

We later show that martingale-admissible kernels lead to martingale prices. Such kernels anticipate the market order flow to set martingale prices according to linear permanent cross-impact. This prevents agents who successfully forecast order flow to trade profitably. However, it does not forbid statistical arbitrages entirely. Before introducing relevant definitions, we define trading strategies within our market model below.

**Definition 1.2** (Trading strategy). *The buy and sell trades sent under the trading strategy  $f: \mathbb{R} \rightarrow \mathbb{R}^d$  are  $d$ -dimensional Poisson processes  $n^a$  and  $n^b$ , independent of the Hawkes process  $(N^a, N^b)$ , with intensities given by  $f^a := \max(f, 0)$  and  $f^b := \max(-f, 0)$ . The (average) cost of the trading strategy  $f$  is*

$$C(f) := \int_0^\infty \int_0^t f(t)^\top K(t-s) f(s) ds dt. \quad (4)$$

If  $\int_0^\infty f(s)ds = 0$ , the trading strategy is called a round-trip strategy and if  $f$  has finite support, it is called a finite-horizon trading strategy.

All trading strategies considered in this paper have deterministic intensity, ignore exchange fees, bid-ask spreads and other microstructural trading costs, so that trading costs are exclusively induced by market impact. The average cost Equation (4) is derived since under the agent's trading strategy the price process becomes

$$P_t = P_0 + \int_0^t K(t-s)(dN_s^a - dN_s^b + dn_s^a - dn_s^b),$$

so that the average trading cost of the strategy is

$$\begin{aligned} \mathbb{E}\left[\int_0^\infty (dn_t^a - dn_t^b)^\top (P_t - P_0)\right] &= \mathbb{E}\left[\int_0^\infty (f^a(t) - f^b(t))^\top (P_t - P_0)dt\right] \\ &= \mathbb{E}\left[\int_0^\infty \int_0^t (f^a(t) - f^b(t))^\top K(t-s)(dN_s^a - dN_s^b + dn_s^a - dn_s^b)dt\right]. \end{aligned}$$

Therefore, since the counting processes  $n^a$  and  $n^b$  are independent from each other and from the Hawkes process  $(N^a, N^b)$  and  $\mathbb{E}[\lambda_t^a] = \mathbb{E}[\lambda_t^b]$  for all  $t$ , we obtain

$$\begin{aligned} \mathbb{E}\left[\int_0^\infty (dn_t^a - dn_t^b)^\top (P_t - P_0)\right] &= \int_0^\infty \int_0^t (f^a(t) - f^b(t))^\top K(t-s)(f^a(s) - f^b(s))dsdt \\ &= \int_0^\infty \int_0^t f(t)^\top K(t-s)f(s)dsdt. \end{aligned}$$

This justifies Equation (4). Trading strategies which are profitable on average are called statistical arbitrages and defined below.

**Definition 1.3** (Statistical arbitrage). *A statistical arbitrage is a finite horizon, round-trip trading strategy such that its costs are negative:*

$$C(f) < 0.$$

Cross-impact kernels which allow for statistical arbitrage induce important issues for applications. For example, they would bias trading strategies which seek to minimise trading costs towards trading-induced price manipulation. For such problems, we require trading costs models with theoretical guarantees of no statistical arbitrage. We call *no-statistical-arbitrage-admissible* (or *nsa-admissible* for short) cross-impact kernels that prevent statistical arbitrage.

**Definition 1.4** (No-statistical-arbitrage-admissible kernels). *A cross-impact kernel  $K$  is said to be nsa-admissible if there are no possible statistical arbitrages, i.e. no round-trip trading strategies with average negative cost.*

Finally, we make the distinction between readily available information, such as prices and trades, and non-directly observable information, such as the cross-impact kernel  $K$  or the permanent cross-impact matrix  $\Lambda$ . We refer to empirical observables for information which is easily measurable.

**Definition 1.5** (Empirical observables). *An empirical observable is a first or second-order moment measure of the price or order flows counting processes.*

Definition of the moment measures are given in Appendix A. Loosely speaking, they can be seen as the moments of our market variables. Empirical observables play an important role: they can be understood as key

features of our stylised market, which we measure and use to derive the cross-impact kernel  $K$ . The next section shows that they constrain the class of relevant cross-impact kernels.

Our model being set, the next section presents the main results of the paper.

## 2 Characterisation of cross-impact kernels

The previous section introduced the framework in force throughout the paper. We now characterise the cross-impact kernels  $K$  which emerge from these assumptions, depending on the hypotheses on the market. We characterise martingale-admissible kernels in Section 2.1 and nsa-admissible kernels in Section 2.2. Finally, Section 2.3 concludes on the cross-impact kernels which are both martingale-admissible and nsa-admissible.

### 2.1 Characterisation of martingale-admissible kernels

In this section, we focus on characterising martingale-admissible kernels. We begin by characterising martingale-admissible cross-impact kernels as a function of Hawkes parameters and the permanent cross-impact matrix  $\Lambda$ . Then, we express these cross-impact kernels as a function of empirical observables.

#### 2.1.1 Cross-impact kernel as a function of Hawkes parameters

The following proposition derives the martingale-admissible cross-impact kernels  $K$  as a function of Hawkes parameters and the permanent cross-impact matrix  $\Lambda$ .

**Proposition 2.1.** *For any martingale-admissible kernel  $K$ , the price is a martingale,  $\Phi^{b/b} - \Phi^{a/b} = \Phi^{a/a} - \Phi^{b/a}$  and for all  $t \in \mathbb{R}_+$*

$$K(t) = K(0)(\mathbb{I}_d - \int_0^t \varphi(s)ds),$$

where we have introduced the imbalance kernel

$$\varphi := \Phi^{b/b} - \Phi^{a/b} = \Phi^{a/a} - \Phi^{b/a}.$$

Furthermore, the immediate cross-impact matrix and permanent cross-impact matrix are related as follows:

$$K(0) = \Lambda \left( \mathbb{I}_d - \int_0^\infty \varphi(s)ds \right)^{-1} \text{diag}(\nu)^{-1}.$$

The proof of Proposition 2.1 is given in Appendix C.1. Note that, by Proposition 2.1 and Assumption 4, any martingale-admissible cross-impact kernel  $K$  is almost-everywhere differentiable and its derivative is square-integrable. Proposition 2.1 provides an expression for  $K$  as a function of Hawkes parameters and the permanent cross-impact matrix  $\Lambda$ . However, the imbalance kernel  $\varphi$  is hard to estimate and we cannot measure the permanent cross-impact matrix  $\Lambda$  on real data. We would like to derive an analogous expression using solely empirical observables which can be easily measured on empirical data. This is the topic of the next section.

### 2.1.2 Cross-impact kernel as a function of empirical observables

To derive an expression for martingale-admissible cross-impact kernels, it is convenient to introduce the stationary version of the Hawkes order flow process. We write  $\tilde{N}$  for the stationary version of the Hawkes process with baseline  $\mu$  and kernel  $\Phi$  (this process exists and is unique since the spectral radius of the  $L_1$  norm of  $\Phi$  is smaller than one, see e.g. [7]). Loosely speaking,  $\tilde{N}$  describes the long-term behaviour of  $N$ . The stationary version allows us to define properly empirical observables but we still consider that the order flow process is given by the non-stationary process.

We write  $\Omega^{\tilde{N}}$  and  $\Omega$  for the reduced covariance measures, defined in Appendix A, of the multivariate stationary point processes  $\tilde{N}$  and  $\tilde{N}^b - \tilde{N}^a$ . By construction, we have  $\Omega = (\mathbb{I}_d, -\mathbb{I}_d)\Omega^{\tilde{N}}(\mathbb{I}_d, -\mathbb{I}_d)^\top$ .

Using the above notations, the following proposition relates martingale-admissible kernels to empirical observables and the boundary values of the cross-impact kernel  $K(0)$  and  $\lim_{t \rightarrow \infty} K(t) = \Lambda$ .

**Proposition 2.2.** *Any martingale-admissible cross-impact kernel  $K$  satisfies, for almost all  $\omega \in \mathbb{R}$*

$$\widehat{K}'(\omega) = \frac{1}{\sqrt{2}} \mathcal{G} \mathcal{O} \mathcal{L}(\omega)^{-1} - K(0), \quad (5)$$

where  $\mathcal{G}$  is any matrix such that  $\mathcal{G}\mathcal{G}^\top = \Lambda \int_0^\infty \Omega(ds) \Lambda^\top$ ,  $\mathcal{L}$  is any spectral factor of  $\Omega$  (see Definition B.2) and  $\mathcal{O}$  is the unique rotation matrix such that

$$\Lambda = \frac{1}{\sqrt{2}} \mathcal{G} \mathcal{O} \mathcal{L}(0)^{-1},$$

where we recall that  $\lim_{t \rightarrow \infty} K(t) = \Lambda$  and, by Proposition 2.1 and Assumption 4,  $K$  is almost-everywhere differentiable and its derivative is square-integrable.

The proof of Proposition 2.2 is given in Appendix C.2. Proposition 2.2 completely characterises the derivative of martingale-admissible cross-impact kernels as a function of quantities easily measurable on data through  $\Sigma, \Omega$  and the boundary values  $K(0)$  and  $\Lambda$ . However, these are not known a priori. Thus, for a given set of empirical observables, martingale-admissible cross-impact kernels may only differ by their boundary values.

The above characterisation for martingale-admissible cross-impact kernels is useful to calibrate martingale-admissible kernels and we make use of it in Section 3. However, to do so, we must choose values for  $K(0)$  and  $\Lambda$ . An outstanding question is thus that of appropriate values, which we address in the next section where we find that nsa-admissible kernels have constrained boundary values.

## 2.2 Characterisation of nsa-admissible kernels

The previous section examined martingale-admissible cross-impact kernels. We found that such cross-impact kernels are completely constrained – except at the boundaries. In this section, we focus on nsa-admissible cross-impact kernels. Contrary to the previous section, we will find that nsa-admissible cross-impact kernels are largely unconstrained, except at the boundary values. We begin by showing the latter.

### 2.2.1 Constraints on the boundary values of the cross-impact kernel

This section derives the boundary values for nsa-admissible cross-impact kernels. The following proposition characterises the immediate cross-impact matrix for any nsa-admissible kernel. We introduce  $\theta := (\mathbb{I}_d - \Phi^{a/a})\mu + \Phi^{a/b}\mu$  which represents the stationary average of the intensity of incoming buy or sell orders.

**Proposition 2.3.** *For any nsa-admissible kernel  $K$ , we have*

$$K(0) = \frac{1}{\sqrt{2}}(\mathcal{L}_0^{-1})^\top \sqrt{\mathcal{L}_0^\top \Sigma \mathcal{L}_0} \mathcal{L}_0^{-1}, \quad (6)$$

where

- (i) *the matrix  $\mathcal{L}_0$  is any matrix such that  $\mathcal{L}_0 \mathcal{L}_0^\top = \text{diag}(\theta_1 v_1^2, \dots, \theta_d v_d^2)$ ,*
- (ii) *the matrix  $\Sigma := \lim_{t \rightarrow \infty} \mathbb{E}[\mathbf{d}\langle P, P \rangle_t]$  is loosely speaking the stationary instantaneous covariance matrix of returns. The existence of this limit is shown in the proof of the proposition.*

The proof of Proposition 2.3 is given in Appendix C.3. Note that  $K(0)$  does not depend on the choice of  $\mathcal{L}_0$  such that  $\mathcal{L}_0 \mathcal{L}_0^\top = \text{diag}(\theta_1 v_1^2, \dots, \theta_d v_d^2)$ . The matrix  $K(0)$  has a microscopic interpretation. Indeed, within Kyle's insider trading model [16] extended to multiple assets [9] the market-maker adjusts his quotes according to the pricing rule  $Gq$ , where  $G$  is called the Kyle cross-impact matrix and  $q = q_{IT} + q_{NT}$  is the aggregate order flow of the insider and noise traders. In this model, if the price-covariance matrix is  $\Sigma$  and the noise order-flow covariance matrix  $\mathbb{E}[(q_{NT})^\top q_{NT}]$  is  $\Omega(\{0\}) = \text{diag}(\theta_1 v_1^2, \dots, \theta_d v_d^2)$ , then  $G = K(0)$ , where  $K(0)$  is given by Proposition 2.3. In our model,  $\Omega(\{0\})$  represents the instantaneous covariance matrix of order flow, which is diagonal since there are no simultaneous orders on different assets. Thus, although no agents in our model have information,  $K(0)$  can be interpreted through the lens of information revelation.

With a quite similar proof as that of Proposition 2.3, we can show that the permanent cross-impact matrix for any nsa-admissible kernel is symmetric non-negative.

**Proposition 2.4.** *For any nsa-admissible kernel  $K$ , the matrix  $\lim_{t \rightarrow \infty} K(t) = \Lambda$  is symmetric non-negative.*

The proof of Proposition 2.4 is given in Appendix C.4. Importantly, the elements of Equation (6) can be estimated quite easily on data. Thus, for any nsa-admissible kernel  $K$ ,  $K(0)$  can be expressed solely as a function of market observables. On the other hand, the permanent cross-impact matrix  $\lim_{t \rightarrow \infty} K(t) = \Lambda$  is only constrained by symmetry and non-negativeness.

### 2.2.2 Constraints on the Fourier transform of the cross-impact kernel

The previous section showed that boundary values of nsa-admissible kernels are constrained and that their boundary value at zero is completely characterised. In this section, we derive necessary and sufficient conditions for nsa-admissible kernels. We begin with a result from [1], which holds in a more general setting than the one in force in this paper, given in the lemma below.

**Lemma 2.1** (Theorem 2.10 of [1]). *A continuous cross-impact kernel  $K$  is nsa-admissible if and only if there exists a matrix-valued non-negative definite Hermitian measure  $\mathbb{M}$  such that for all  $t \in \mathbb{R}$  we have*

$$\mathcal{Z}(t) = \int_{\mathbb{R}} e^{i\gamma t} \mathbb{M}(d\gamma),$$

where

$$\mathcal{Z}(t) := \begin{cases} K(t) & \text{if } t > 0 \\ K(0) & \text{if } t = 0 \\ K(-t)^\top & \text{if } t < 0 \end{cases}.$$

*Proof.* Since  $K(0)$  is symmetric and  $K$  is continuous this result stems from Theorem 2.10 of [1].  $\square$

In our framework, we can extend the previous result and prove some properties concerning the smoothness of nsa-admissible kernels, which is the topic of the next proposition.

**Proposition 2.5.** *A continuous cross-impact kernel  $K$  is nsa-admissible kernel if and only if, using the notations of Lemma 2.1, one of the following identities holds for almost all  $t \in \mathbb{R}$*

$$\begin{aligned} \mathcal{Z}'(t) &= i \int_{\mathbb{R}} \gamma e^{i\gamma t} \mathbb{M}(\mathrm{d}\gamma) \\ \mathcal{Z}''(t) &= - \int_{\mathbb{R}} \gamma^2 e^{i\gamma t} \mathbb{M}(\mathrm{d}\gamma), \end{aligned}$$

where  $\mathbb{M}$  is a matrix-valued non-negative definite Hermitian measure such that  $\mathcal{Z}(t) = \int_{\mathbb{R}} e^{i\gamma t} \mathbb{M}(\mathrm{d}\gamma)$  and each integral converges absolutely, i.e. each integrand is absolutely integrable with respect to the measure  $\mathbb{M}$ . If any of these conditions is satisfied, the matrix function  $\mathcal{Z}$  has twice continuously differentiable entries.

The proof of Proposition 2.5 is given in Appendix C.5. The regularity properties of Proposition 2.5 enable us to check monotonicity and convexity of continuous nsa-admissible cross-impact kernels, a topic of interest as discussed in [1].

### 2.3 Characterisation of martingale and nsa-admissible kernels

This section summarises the results from the two previous ones to characterise kernels which are martingale-admissible and nsa-admissible. We have seen that martingale-admissible kernels are constrained everywhere except at the boundaries, while nsa-admissible kernels are constrained at the boundaries but are largely unconstrained elsewhere. It is thus natural to observe that there is at most one kernel that is both martingale-admissible and nsa-admissible. In practice, this candidate kernel is not always nsa-admissible. For applications, it may be interesting to slightly relax the martingale property in order to guarantee no statistical arbitrage. Thus, we introduce a regularisation technique to obtain a kernel close to the kernel which gives martingale prices but that prevents arbitrage.

The next proposition shows that for kernels which are both nsa-admissible and martingale-admissible, the permanent cross-impact matrix is fixed. We recall that the matrix  $\Sigma = \lim_{t \rightarrow \infty} \mathbb{E}[\mathrm{d}\langle P, P \rangle_t]$  is loosely speaking the stationary instantaneous covariance matrix of returns.

**Proposition 2.6.** *For any nsa-admissible, martingale-admissible kernel  $K$ , we have*

$$\lim_{t \rightarrow \infty} K(t) = \Lambda = \frac{1}{\sqrt{2}} (\mathcal{L}_\infty^{-1})^\top \sqrt{\mathcal{L}_\infty^\top \Sigma \mathcal{L}_\infty} \mathcal{L}_\infty^{-1}, \quad (7)$$

where the matrix  $\mathcal{L}_\infty$  is any matrix such that  $\mathcal{L}_\infty \mathcal{L}_\infty^\top = \int_0^\infty \Omega(\mathrm{d}s)$ , which is loosely speaking the stationary total autocovariance matrix of order flows.

*Proof.* As  $K$  is both martingale-admissible and nsa-admissible, Propositions 2.2 to 2.4 imply that  $\Lambda$  is a symmetric, non-negative matrix that satisfies

$$\Lambda \int_0^\infty \Omega(ds) \Lambda^\top = \frac{1}{2} \Sigma.$$

The result follows.  $\square$

The previous proposition highlights that, as  $K(0)$  in Proposition 2.3, the permanent cross-impact matrix given in Equation (7) has a microscopic interpretation. However, while  $K(0)$  can be interpreted as the market-maker pricing rule in a market where the price covariance is  $\Sigma$  and the noise order flow covariance matrix is  $\Omega(\{0\}) = \text{diag}(\theta_1 v_1^2, \dots, \theta_d v_d^2)$ , the permanent cross-impact matrix  $\Lambda$  can be interpreted as the market-maker pricing rule in a market where the price covariance is  $\Sigma$  and the order flow covariance is  $\int_0^\infty \Omega(ds)$ . The latter can be interpreted as the total order flow covariance which encapsulates instantaneous and non-instantaneous liquidity.

The next theorem summarises the results of the paper to completely characterise cross-impact kernels which are both martingale-admissible and nsa-admissible.

**Theorem 2.1.** *There exists a unique cross-impact kernel  $K$  that is martingale-admissible and which satisfies the necessary conditions for arbitrage-admissibility outlined in Equations (6) and (7). Its expression is given by inverting Equation (5) and setting the boundary values given by Equations (6) and (7). Furthermore, if it satisfies Proposition 2.5, then  $K$  is also nsa-admissible.*

*Proof.* Let  $K$  be a cross-impact kernel which is both martingale-admissible and satisfies Equations (6) and (7). Then by Proposition 2.1, it is continuous. Furthermore, it must satisfy Propositions 2.1 and 2.3 so that it is unique and its boundary values are given by Equations (6) and (7). Since it is continuous, it is nsa-admissible if and only if it satisfies the necessary and sufficient conditions of Proposition 2.5 (or Lemma 2.1).  $\square$

The theorem shows there exists only one martingale-admissible kernel which satisfies the boundary conditions of nsa-admissible kernels. Though it may not be nsa-admissible, it is certainly closer to being nsa-admissible than other martingale-admissible kernels since it satisfies necessary conditions of nsa-admissible kernels. Given its importance, we write this cross-impact kernel  $K^1$  in the following.

**Definition 2.1** ( $K^1$  kernel). *The cross-impact kernel  $K^1$  is the unique martingale-admissible kernel that satisfies Equations (6) and (7).*

While  $K^1$  is a good candidate for applications, we have no guarantee that this kernel is nsa-admissible. This naturally poses issues in certain applications. For example, in portfolio optimization, a trading cost model which allows for arbitrages induces spurious round-trip strategies, as shown in [1]. Thus, we introduce a regularisation method to find a kernel close to this candidate but which is nsa-admissible, which we write  $K^2$ . This motivates to the following definition.

**Definition 2.2** ( $K^2$  kernel). *The cross-impact kernel  $K^2$  is defined as*

$$K^2 = \underset{\text{Karbitrage-admissible}}{\operatorname{argmin}} \left\| \widehat{K}^1 - \widehat{K} \right\|_F,$$

where  $\|\cdot\|_F$  is the Frobenius norm.

The kernel  $K^2$  exists and is unique so that the previous definition is justified. Indeed, by Lemma 2.1 and Proposition 2.5, any nsa-admissible kernel  $K$  must be such that  $\hat{K} + \hat{K}^*$  is non-negative. Therefore

$$\operatorname{argmin}_{K \text{arbitrage-admissible}} \|\hat{K}^1 - \hat{K}\|_F = \operatorname{argmin}_{\hat{K} + \hat{K}^* \geq 0} \|\hat{K}^1 - \hat{K}\|_F.$$

Thus,  $K^2$  can be computed from  $K^1$  by replacing each eigenvalue  $\rho$  of  $\hat{K}^1 + \hat{K}^{1*}$  by  $\max(\rho, 0)$ . Loosely speaking,  $K^2$  trades martingale-admissibility for arbitrage-admissibility while staying close to  $K^1$ .

The next section applies our results to market data to compute  $K^1$  and  $K^2$ .

### 3 Application to financial data

This section focuses on applying the previous results to compute the kernels  $K^1$  and  $K^2$  on market data. Details on the methodology and data used are given in Appendix D.

The dataset used comprises of 4 years of volumes and price data for two maturities of E-Mini SP500 futures traded on the CME, so that  $d = 2$ . These futures are financially settled at expiry (in addition to daily settlements) according to the value of the SP500 index. The two futures selected are the lead month future, referred to as SPMINI, and the third upcoming future, referred to as SPMINI3. This data has been explored in a previous study and we refer the reader to [22] for more details into the underlying data and processing methodology.<sup>1</sup>

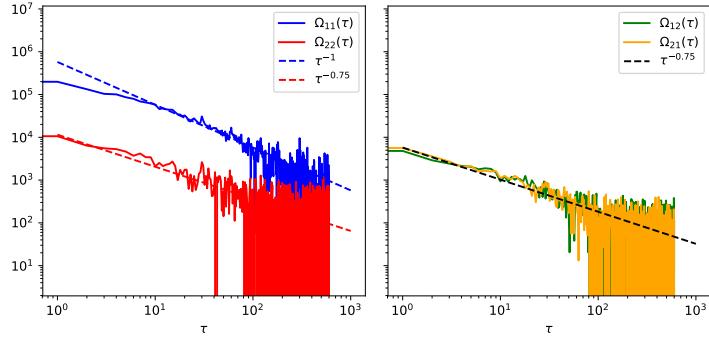
We begin by reporting the relevant empirical observable of our system, namely the reduced covariance measure  $\Omega$  and the price-covariance matrix  $\Sigma$ .

Figure 1 reports empirical estimates of  $\Omega([\tau\tau_s, (\tau+1)\tau_s])$  for different values of  $\tau$ , for a time resolution of  $\tau_s = 1$  second. By a slight abuse of notation, we write  $\Omega(\tau)$  for  $\Omega([\tau\tau_s, (\tau+1)\tau_s])$ . With these conventions,  $\Omega(0)$  represents the trade covariance, so that  $\Omega_{11}(0)$  is the instantaneous variance of signed order flow on the front month future SPMINI and  $\Omega_{22}(0)$  the instantaneous variance of signed order flow on SPMINI3. These quantities reflect the liquidity of the underlying assets since they increase with the daily traded volume [22]. The figure shows that the front-month maturity SPMINI is approximately 10 times more liquid than the SPMINI3, which highlights that most trading occurs on the leading month contract. The order flow auto-covariances  $\Omega_{11}(\tau)$  and  $\Omega_{22}(\tau)$  are slowly decaying in  $\tau$ , although  $\Omega_{11}(\tau)$  exhibits faster decay than  $\Omega_{22}(\tau)$ . Furthermore, we observe that  $\Omega_{12}(\tau) \approx \Omega_{21}(\tau)$  so that there are no lead-lag effects in the order flows. This shows that  $\Omega$  cannot be easily factorized under the form  $\Omega(\tau) \approx \tau^{-\beta} C$ , with some  $\beta < 1$  and  $C \in \mathcal{M}_d(\mathbb{R})$ . This hypothesis is used in certain cross-impact kernels [5].

The price-covariance matrix  $\Sigma$ , not shown here but reported in [22], shows strong correlation ( $\approx 90\%$ ) between the two maturities. This is natural since both futures have the same underlying.

---

<sup>1</sup>The authors thank the *Econophysics & Complex Systems* Research Chair for providing access to this data.

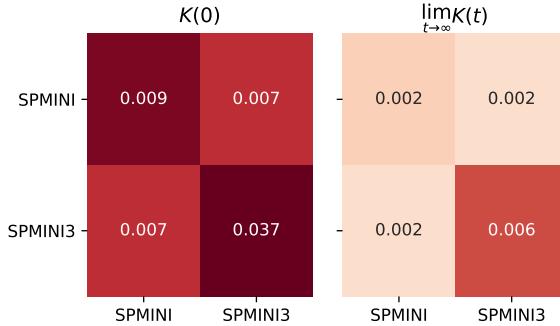


**Figure 1: Order flow auto-covariance  $\Omega$ .**

Diagonal elements of the auto-covariance measure (left) and off-diagonal elements of the auto-covariance measure (right). Dashed lines represent power-law fits of the auto-covariances. Order flow is in product units: number of contracts traded times the contract valpoint (which is 50 for these futures). Details about the estimation procedure are given in Appendix D.2.

Figure 2 shows the estimated boundary values of the  $K^1$  and  $K^2$  kernels which are by definition the same for  $K^1$  and  $K^2$ , computed using Propositions 2.3 and 2.6. The calibrated values confirm some intuitive ideas:

- price impact on the liquid future is lower than on the illiquid future:  $K_{11}(0) < K_{22}(0)$  and  $\Lambda_{11} < \Lambda_{22}$ ;
- buying a future immediately pushes the price of the other, as  $K_{12}(0), K_{21}(0) > 0$ , and the permanent impact contribution is positive since  $\Lambda_{12}, \Lambda_{21} > 0$ ;
- permanent impact is lower than immediate impact, as each component of the immediate impact matrix is larger than the permanent cross-impact matrix:  $K^1(0) = K^2(0) > \Lambda$ .

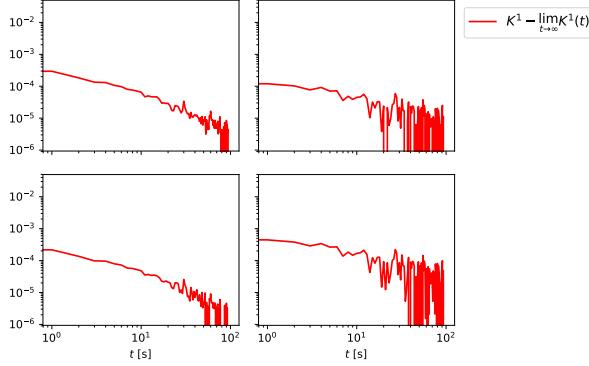


**Figure 2: Boundary values of  $K^1$  and  $K^2$ .**

The boundary values of the martingale-admissible kernel  $K^1$  and the nsa-admissible kernel  $K^2$ , the immediate cross-impact matrix  $K^1(0) = K^2(0)$  and the permanent cross-impact matrix  $\Lambda = \lim_{t \rightarrow \infty} K^1(t) = \lim_{t \rightarrow \infty} K^2(t)$ , estimated using Proposition 2.3 and Proposition 2.6. Values are reported in basis points (i.e.  $10^4$  of their units). Details about the estimation procedure are given in Appendix D.4.

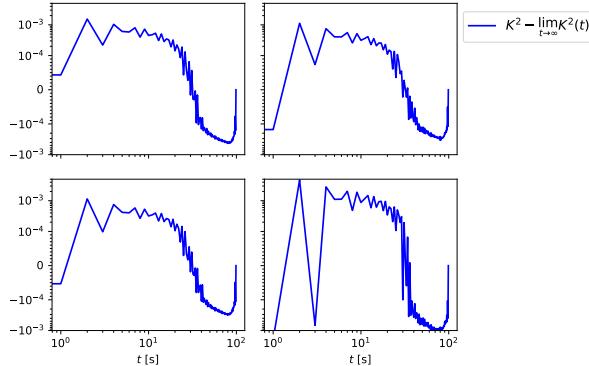
Figures 3 and 4 show both cross-impact kernels. The boundary values of  $K^1$  and  $K^2$  are identical and given in Figure 2. As could be checked numerically, the martingale-admissible kernel  $K^1$  is not nsa-admissible. Thus, there is no kernel which is both martingale-admissible and nsa-admissible here. Given the very strong correlations between both assets, it is not surprising that  $K_{21}^2 \approx K_{11}^2$  and  $K_{21}^1 \approx K_{11}^1$ : trading the first maturity

pushes the price of each future by roughly the same amount. Finally, a somewhat surprising feature of  $K^2$  is that it is non-monotonous. The cross-impact kernel  $K^2$  is sensitive to numerical errors in the estimation methodology detailed in Appendix D.4, which may explain the strange value of a point of  $K_{22}^2$ . Nevertheless, it has little incidence on the kernel's fit to data (as we can see from Figure 5).



**Figure 3: Values of the  $K^1$  kernel.**

The values of the transient part of the martingale-admissible  $K^1 - \lim_{t \rightarrow \infty} K^1(t) = K^1 - \Lambda$  (red) are reported. Each subplot shows the matrix elements of the kernels. For instance, the top left plot shows  $K^1_{11} - \Lambda_{11}$  and the top right shows  $K^1_{12} - \Lambda_{12}$ . The permanent cross-impact matrix  $\Lambda$  has been removed to highlight the power-law decay of the cross-impact kernel  $K^1$  toward its limit.

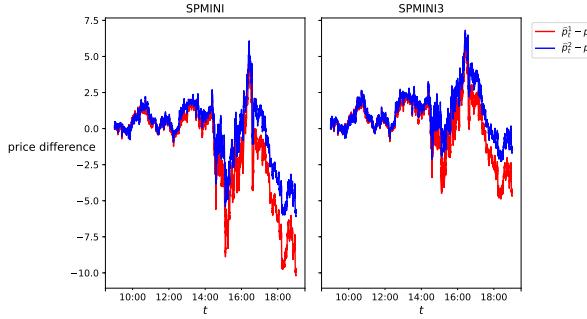


**Figure 4: Values of the  $K^2$  kernel.**

The values of the transient part of the nsa-admissible kernel  $K^2 - \lim_{t \rightarrow \infty} K^2(t) = K^2 - \Lambda$  (blue) are reported. Each subplot shows the matrix elements of the kernels. For instance, the top left plot shows  $K^2_{11}$  and  $K^2_{11} - \Lambda_{11}$  and the top right shows  $K^2_{12} - \Lambda_{12}$ . The permanent cross-impact matrix  $\Lambda$  has been removed to highlight the behaviour of the transient part of the cross-impact kernel  $K^2$ .

We illustrate the predictions of the different kernels in Figure 5. For a given trading day taken on the 31st of January 2017, we measure the traded order flows and build the predicted price changes according to the cross-impact rule Equation (2). We complete this procedure with the martingale-admissible kernel  $K^1$  and the nsa-admissible kernel  $K^2$ . We then compare the predicted price changes to the actual price changes. As is consistent with the literature on price impact, we see that price changes predicted from our cross-impact models are qualitatively consistent with realised price changes. However, cross-impact models are able to use trades on the liquid maturity (SPMINI) to explain price changes on the illiquid maturity (SPMINI3). This is

critical since the prices of the two maturities are strongly correlated but most trades occur on the leading month contract. Although the two kernels are quite different, their predictions are strikingly similar. This highlights that our regularisation procedure was successful in finding an nsa-admissible kernel that fits data well and prevents statistical arbitrage.



**Figure 5: Example of predicted prices from  $K^1$  and  $K^2$ .**

The difference between predicted prices between the realized price and the impact-induced price from Equation (1) with the martingale-admissible cross-impact kernel  $K^1$ , given by  $\hat{p}_t^1 = p_0 + \sum_{s < t} K^1(t-s)q_s$ , (red) and with the nsa-admissible cross-impact kernel  $K^2$ , given by  $\hat{p}_t^2 = p_0 + \sum_{s < t} K^2(t-s)q_s$ , (blue), where  $(q_t)$  are measured signed volumes of market orders. Price difference is reported in dollars per lot.

## 4 Conclusion

The goal of this paper was to characterise the class of cross-impact kernels which anticipate order flows and lead to martingale prices (*martingale-admissible* kernels) and the class of cross-impact kernels which prevent statistical arbitrage (*no-statistical-arbitrage-admissible* kernels). To do so, we introduce a market where trades are self-exciting and price impact is linear in the total market order flow. In this model, we derive necessary and sufficient conditions for nsa-admissible and martingale-admissible kernels. In particular, we show that only one candidate, dubbed the  $K^1$  kernel, could be both martingale-admissible and nsa-admissible. As there are no theoretical guarantees that the  $K^1$  kernel prevents arbitrage, we introduce the  $K^2$  kernel which is close to  $K^1$  and prevents arbitrage. We find formulas for calibration of both kernels and apply them on SP500 futures.

One key result of this paper is that, given a set of market conditions, namely the asset price covariance which encodes co-movement of assets and the auto-covariance of trades which captures the way agents trade in the market, it may not be possible to conciliate the two notions of price efficiency: martingale prices and no-arbitrage. This is specific to the multi-asset case since in the single-asset case the two notions nicely co-exist [14, 15]. Therefore, a problem we hope to address in future work is to find non-trivial market conditions where there exists a cross-impact kernel which is both nsa-admissible and martingale-admissible.

Finally, the results outlined in the paper have applications in market-making and trading costs estimation. For market-makers, the cross-impact kernels can be calibrated in practice to better capture adverse selection or price decay after trades. For portfolio managers, these models provide better estimate of trading costs and could be used to derive more optimal trading strategies which account for cross-impact and its decay.

## References

- [1] Aurélien Alfonsi, Florian Klöck, and Alexander Schied. Multivariate transient price impact and matrix-valued positive definite functions. *Mathematics of operations research*, 41(3):914–934, 2016.
- [2] Robert Almgren, Chee Thum, Emmanuel Hauptmann, and Hong Li. Direct estimation of equity market impact. *Risk*, 18(7):58–62, 2005.
- [3] Emmanuel Bacry, Khalil Dayri, and Jean-François Muzy. Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *The European Physical Journal B*, 85(5):1–12, 2012.
- [4] Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.
- [5] Michael Benzaquen, Iacopo Mastromatteo, Zoltan Eisler, and Jean-Philippe Bouchaud. Dissecting cross-impact on stock markets: An empirical analysis. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(2):23406, 2017.
- [6] Jean-Philippe Bouchaud, Julius Bonart, Jonathan Donier, and Martin Gould. *Trades, Quotes and Prices*. Cambridge University Press, 2018.
- [7] Pierre Brémaud and Laurent Massoulié. Stability of nonlinear Hawkes processes. *The Annals of Probability*, pages 1563–1588, 1996.
- [8] Daryl J Daley and David Vere-Jones. *An introduction to the theory of point processes: volume II: general theory and structure*. Springer New York, 2008.
- [9] Luis Carlos Garcia del Molino, Iacopo Mastromatteo, Michael Benzaquen, and Jean-Philippe Bouchaud. The multivariate Kyle model: More is different. *SIAM Journal on Financial Mathematics*, 11(2):327–357, 2020.
- [10] Jim Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759, 2010.
- [11] Stephen J Hardiman, Nicolas Bercot, and Jean-Philippe Bouchaud. Critical reflexivity in financial markets: a hawkes process analysis. *The European Physical Journal B*, 86(10):1–9, 2013.
- [12] Joel Hasbrouck and Duane J Seppi. Common factors in prices, order flows, and liquidity. *Journal of financial Economics*, 59(3):383–411, 2001.
- [13] Alan. G Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(3):438–443, 1971.
- [14] Thibault Jaisson. Market impact as anticipation of the order flow imbalance. *Quantitative Finance*, 15(7):1123–1135, 2015.
- [15] Paul Jusselin and Mathieu Rosenbaum. No-arbitrage implies power-law market impact and rough volatility. *Mathematical Finance*, 2018.

- [16] Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.
- [17] Eugene Lukacs. *Characteristic functions*. Griffin, 1970.
- [18] John G McWhirter, Paul D Baxter, Tom Cooper, Soydan Redif, and Joanne Foster. An EVD algorithm for para-hermitian polynomial matrices. *IEEE Transactions on Signal Processing*, 55(5):2158–2169, 2007.
- [19] Paolo Pasquariello and Clara Vega. Strategic cross-trading in the US stock market. *Review of Finance*, 19(1):229–282, 2015.
- [20] M. Schneider and F. Lillo. Cross-impact and no-dynamic-arbitrage. *Quantitative Finance*, 19(1):137–154, 2019.
- [21] Edward C Titchmarsh. *Introduction to the theory of Fourier integrals*. Clarendon Press, 1948.
- [22] Mehdi Tomas, Iacopo Mastromatteo, and Michael Benzaquen. How to build a cross-impact model from first principles: Theoretical requirements and empirical results. *arXiv preprint arXiv:2004.01624*, 2020.
- [23] Mehdi Tomas, Iacopo Mastromatteo, and Michael Benzaquen. Cross impact in derivative markets. *arXiv preprint arXiv:2102.02834*, 2021.
- [24] Mehdi Tomas and Mathieu Rosenbaum. From microscopic price dynamics to multidimensional rough volatility models. *Advances in Applied Probability, to appear*, 2019.
- [25] Nicolo Torre. BARRA market Impact model handbook. *BARRA Inc., Berkeley*, 1997.
- [26] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.
- [27] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.
- [28] Shanshan Wang, Sebastian Neustüß, and Thomas Guhr. Grasping asymmetric information in market impacts. *arXiv preprint arXiv:1710.07959*, 2017.
- [29] Shanshan Wang, Rudi Schäfer, and Thomas Guhr. Price response in correlated financial markets: empirical results. *arXiv preprint arXiv:1510.03205*, 2015.
- [30] Shanshan Wang, Rudi Schäfer, and Thomas Guhr. Cross-response in correlated financial markets: individual stocks. *The European Physical Journal B*, 89(4):105, 2016.
- [31] Zeliang Wang, John G McWhirter, and Stephan Weiss. Multichannel spectral factorization algorithm using polynomial matrix eigenvalue decomposition. In *2015 49th Asilomar conference on signals, systems and computers*, pages 1714–1718. IEEE, 2015.

- [32] Norbert Wiener and E.J. Akutowicz. A factorization of positive hermitian matrices. *Journal of Mathematics and Mechanics*, pages 111–120, 1959.
- [33] Norbert Wiener and Pesi Masani. The prediction theory of multivariate stochastic processes. *Acta Mathematica*, 98(1-4):111–150, 1957.

## A Moment measures for point processes

This section presents definitions of moment measures for point processes. Throughout this section,  $\mathbb{X}$  denotes some Borelian subset of  $\mathbb{R}^d$  and  $\mathcal{X} = \mathcal{B}(\mathbb{X})$ .

**Definition A.1** ( $n$ -th moment measure). *Given a univariate stationary point process  $N$  on  $\mathbb{X}$ , the first and second order moment measure of the point process  $N$  are measures on  $\mathbb{X}$  and  $\mathbb{X}^2$  defined for all  $A, B \in \mathcal{X}$  as*

$$M_1(A) = \mathbb{E}[N(A)]$$

$$M_2(A \times B) = \mathbb{E}[N(A)N(B)]$$

whenever these expectations exist.

A key quantity is the reduced measure second order measure which we introduce below.

**Definition A.2** (Reduced measure (Proposition 12.6.III of [8])). *There exist a reduced measure, noted  $\overset{\vee}{M}_2$ , such that for any bounded measurable function  $f$  on  $\mathbb{X}^2$  we have*

$$\int_{\mathbb{X}} f(x_1, x_2) M_2(dx_1, dx_2) = \int_{\mathbb{X}} \int_{\mathbb{X}} f(x, x+y), \overset{\vee}{M}_2(dy) dx$$

The existence result of the reduced measure satisfying the above is shown in Proposition 12.6.III of [8]. Of particular interest to us are the second-order reduced covariance measure introduced below.

**Definition A.3** (Reduced covariance measure). *The reduced covariance measure of a stationary point process is defined as*

$$\overset{\vee}{C}_2(du) = \overset{\vee}{M}_2(du) - m^2 l(du),$$

where  $l$  is the Lebesgue measure,  $m$  is the mean intensity i.e the non-negative constant such that  $M_1(du) = ml(du)$ .

Extensions of the above concepts to multivariate point processes are straightforward. In particular we have the following definition.

**Definition A.4** (second order auto-moment measure). *Given a  $k$ -dimensional, stationary point process  $N$  on  $\mathbb{X}$ , the second order auto-moment measure of the point process  $N$  is a measure on  $\mathbb{X}^2$  defined for all  $A, B \in \mathcal{X}$  and  $1 \leq i, j \leq k$  as*

$$M_{ij}(A \times B) = \mathbb{E}[N_i(A)N_j(B)],$$

whenever this expectation exists.

**Definition A.5** (reduced covariance measure). *The reduced covariance measure of a  $k$ -dimensional, stationary point process is defined, for all  $1 \leq i, j \leq k$  as*

$$\overset{\vee}{C}_{ij}(du) = \overset{\vee}{M}_{ij}(du) - m_i m_j l(du),$$

where  $l$  is the Lebesgue measure,  $m_i$  is the mean intensity i.e the non-negative constant such that  $M_i(du) = m_i l(du)$ .

## B Technical results

This section presents some key technical results necessary for the proofs. We begin by introducing an important functional space for our results, the Hardy space  $\mathbb{H}^2$ .

**Definition B.1** (Hardy space  $\mathbb{H}^2$ ). *The Hardy space  $\mathbb{H}^2$  is the space of functions  $F: \mathbb{C} \rightarrow \mathbb{C}$  such that the following conditions are satisfied:*

- (i) *the function  $F$  is holomorphic in the upper half of the complex plane,*
- (ii) *there exists a constant  $C > 0$  such that, for all  $\xi > 0$*

$$\int_{-\infty}^{\infty} |F(\omega + i\xi)|^2 d\omega < C,$$

- (iii) *for almost all  $\omega \in \mathbb{R}$ , we have*

$$\lim_{\xi \rightarrow 0} F(\omega + i\xi) = F(\omega).$$

Functions in the Hardy space  $\mathbb{H}^2$  are closely related to the Fourier transform of causal functions as shown by the so-called Titchmarsh theorem below.

**Theorem B.1** (Titchmarsh theorem [21]). *Given a real, complex-valued square-integrable function  $F$ , the following conditions are equivalent:*

- (i) *the inverse Fourier transform of  $F$  is a causal function,*
- (ii) *the function  $F$  belongs to the Hardy space  $\mathbb{H}^2$ .*

A key theorem used in the paper to derive the form of martingale-admissible cross-impact kernels is the matrix spectral factorization theorem.

**Theorem B.2** (Matrix spectral factorization theorem [32, 33]). *Let  $F: \mathbb{C} \rightarrow \mathcal{M}_d(\mathbb{R})$  be a matrix function such that  $F$  is positive definite almost everywhere on the unit circle  $\mathbb{T} := \{z \in \mathbb{C}: |z| = 1\}$  with integrable entries on the unit circle such that the Paley-Wiener condition*

$$\log \det F \in L_1(\mathbb{T}) \tag{8}$$

*is satisfied. Then  $F$  admits a factorization on  $\mathbb{T}$*

$$F(z) = F_+(z) F_+(z)^*, \tag{9}$$

*where  $F_+$  is an analytic function with entries in the Hardy space  $\mathbb{H}^2$ . Furthermore, the spectral factor  $F_+$  is unique up to right multiplication by a unitary matrix (i.e. a matrix  $U$  such that  $UU^* = \mathbb{I}_d$ ).*

As spectral factors are referenced throughout the paper, we introduce for convenience a definition below.

**Definition B.2** (Spectral factor). *Let  $F: \mathbb{C} \rightarrow \mathcal{M}_d(\mathbb{R})$  be a matrix function which satisfies the hypotheses of Theorem B.2. Then, a spectral factor  $F_+$  of  $F$  is an analytic function  $F_+$  with entries in the Hardy space  $\mathbb{H}^2$  which satisfies Equation (9).*

## C Proofs

### C.1 Proof of Proposition 2.1

The proof is adapted from [14]. Throughout this section we assume that there exists a martingale-admissible kernel  $K$  such that the price satisfies

$$P_t = P_0 + \int_0^t K(t-s) d(N_s^a - N_s^b) = P_0 + \Lambda \lim_{s \rightarrow \infty} \mathbb{E}[V_s^a - V_s^b | \mathcal{F}_t].$$

To compute the right hand side term, we use a classical result on Hawkes processes [4, 14, 15]: for all  $t \geq 0$ , we have

$$\begin{pmatrix} \lambda_t^a \\ \lambda_t^b \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} + \int_0^t \Psi(t-s) \begin{pmatrix} \mu \\ \mu \end{pmatrix} ds + \int_0^t \Psi(t-s) dM_s, \quad (10)$$

where  $\Psi := \sum_{n \geq 1} \Phi^{*n}$ ,  $\Phi^{*n}$  being the  $n$ -th convolution product of the matrix function  $\Phi$  and  $M$  is a martingale. The next lemma shows that the computation of the conditional expectation of the Hawkes process reduces to the computation of the conditional expectation of the intensity.

**Lemma C.1.** *For all  $t, s \geq 0$  such that  $s \geq t$ , we have*

$$\mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] = N_t^a - N_t^b + \int_t^s \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du.$$

*Proof.* Using the martingale decomposition of the Hawkes process, we have  $N_t = M_t + \int_0^t \lambda_s ds$  where  $M$  is a martingale. Therefore,

$$N_s^a - N_s^b = M_s^a - M_s^b + \int_0^t (\lambda_u^a - \lambda_u^b) du + \int_t^s (\lambda_u^a - \lambda_u^b) du.$$

Using the martingale property, we obtain

$$\mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] = N_t^a - N_t^b + \int_t^s \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du.$$

□

The next lemma computes the conditional expectation of the intensity for our Hawkes processes. It generalises Proposition 3.2 of [14].

**Lemma C.2.** *For all  $t, s \geq 0$  such that  $s \geq t$ , we have*

$$\begin{aligned} \int_t^s \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du &= \int_0^t \int_{t-r}^{s-r} \Xi(u) du (dN_r^a - dN_r^b) \\ &\quad - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{a/a}(x) dx du dN_r^a - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{b/a}(x) dx du dN_r^a \\ &\quad + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{a/b}(x) dx du dN_r^b + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{b/b}(x) dx du dN_r^b, \end{aligned}$$

where  $\Xi := \Psi^{a/a} - \Psi^{b/a} = \Psi^{b/b} - \Psi^{a/b}$ .

*Proof.* We use Equation (10) to derive

$$\int_t^s \mathbb{E}[\lambda_u^a | \mathcal{F}_t] du = \mathbb{E} \left[ \int_t^s \int_0^u \Psi^{a/a}(u-x) dM_x^a du | \mathcal{F}_t \right] + \mathbb{E} \left[ \int_t^s \int_0^u \Psi^{a/b}(u-x) dM_x^b du | \mathcal{F}_t \right].$$

Using the fact that  $M$  is a martingale, we have

$$\begin{aligned} \mathbb{E} \left[ \int_t^s \int_0^u \Psi^{a/a}(u-x) dM_x^a du | \mathcal{F}_t \right] &= \int_t^s \int_0^t \Psi^{a/a}(u-x) dM_x^a du \\ &= \int_t^s \int_0^t \Psi^{a/a}(u-x) \left( dN_x^a - (\mu + \int_0^t \Phi^{a/a}(x-r) dN_r^a + \int_0^t \Phi^{a/b}(x-r) dN_r^b) dx \right) du \\ &= \int_t^s \int_0^t \Psi^{a/a}(u-x) dN_x^a du - \int_t^s \int_0^t \Psi^{a/a}(u-x) \mu dx du \\ &\quad - \int_t^s \int_0^t \Psi^{a/a}(u-x) \int_0^t \Phi^{a/a}(x-r) dN_r^a dx du \\ &\quad - \int_t^s \int_0^t \Psi^{a/a}(u-x) \int_0^t \Phi^{a/b}(x-r) dN_r^b dx du. \end{aligned}$$

Consequently,

$$\begin{aligned} \int_t^s \mathbb{E}[\lambda_u^a | \mathcal{F}_t] du &= \int_t^s \int_0^t \Psi^{a/a}(u-x) dN_x^a du + \int_t^s \int_0^t \Psi^{a/b}(u-x) dN_x^b du \\ &\quad - \int_t^s \int_0^t (\Psi^{a/a} + \Psi^{a/b})(u-x) \mu dx du \\ &\quad - \int_t^s \int_0^t \Psi^{a/a}(u-x) \int_0^t \Phi^{a/a}(x-r) dN_r^a dx du - \int_t^s \int_0^t \Psi^{a/b}(u-x) \int_0^t \Phi^{b/a}(x-r) dN_r^a dx du \\ &\quad - \int_t^s \int_0^t \Psi^{a/a}(u-x) \int_0^t \Phi^{a/b}(x-r) dN_r^b dx du - \int_t^s \int_0^t \Psi^{a/b}(u-x) \int_0^t \Phi^{b/b}(x-r) dN_r^b dx du. \end{aligned}$$

Regrouping terms from  $\lambda^a$  and  $\lambda^b$  we obtain

$$\begin{aligned} \int_t^s \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du &= \int_t^s \int_0^t (\Psi^{a/a} - \Psi^{b/a})(u-x) dN_x^a du + \int_t^s \int_0^t (\Psi^{a/b} - \Psi^{b/b})(u-x) dN_x^b du \\ &\quad - \int_t^s \int_0^t (\Psi^{a/a} + \Psi^{a/b} - \Psi^{b/a} - \Psi^{b/b})(u-x) \mu dx du \\ &\quad - \int_t^s \int_0^t (\Psi^{a/a} - \Psi^{b/a})(u-x) \int_0^t \Phi^{a/a}(x-r) dN_r^a dx du \\ &\quad - \int_t^s \int_0^t (\Psi^{a/b} - \Psi^{b/b})(u-x) \int_0^t \Phi^{b/a}(x-r) dN_r^a dx du \\ &\quad - \int_t^s \int_0^t (\Psi^{a/a} - \Psi^{b/a})(u-x) \int_0^t \Phi^{a/b}(x-r) dN_r^b dx du \\ &\quad - \int_t^s \int_0^t (\Psi^{a/b} - \Psi^{b/b})(u-x) \int_0^t \Phi^{b/b}(x-r) dN_r^b dx du. \end{aligned}$$

By Assumption 2, the above must converge to a finite limit as  $s$  tends to infinity. Therefore, as  $\mu \neq 0$  and all elements of the matrix function  $\Phi$  are non-negative we must have for all  $t \geq 0$  that  $\Psi^{a/a}(t) + \Psi^{a/b}(t) =$

$\Psi^{b/a}(t) + \Psi^{b/b}(t)$ . Therefore, writing  $\Xi := \Psi^{a/a} - \Psi^{b/a} = \Psi^{b/b} - \Psi^{a/b}$ , we have

$$\begin{aligned} \int_t^s \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du &= \int_t^s \int_0^t \Xi(u-x)(dN_x^a - dN_x^b) du \\ &\quad - \int_t^s \int_0^t \Xi(u-x) \int_0^t \Phi^{a/a}(x-r) dN_r^a dx du - \int_t^s \int_0^t \Xi(u-x) \int_0^t \Phi^{b/a}(x-r) dN_r^a dx du \\ &\quad + \int_t^s \int_0^t \Xi(u-x) \int_0^t \Phi^{a/b}(x-r) dN_r^b dx du + \int_t^s \int_0^t \Xi(u-x) \int_0^t \Phi^{b/b}(x-r) dN_r^b dx du. \end{aligned}$$

Using the change of variables  $u' = u - r$  and  $x' = x - r$  we obtain

$$\begin{aligned} \int_t^s \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du &= \int_0^t \int_{t-x}^{s-x} \Xi(u') du' (dN_x^a - dN_x^b) \\ &\quad - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u' - x') \Phi^{a/a}(x') dx' du' dN_r^a - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u' - x') \Phi^{b/a}(x') dx' du' dN_r^a \\ &\quad + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u' - x') \Phi^{a/b}(x') dx' du' dN_r^b + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u' - x') \Phi^{b/b}(x') dx' du' dN_r^b. \end{aligned}$$

□

The next lemma uses the previous results to compute the conditional expectation of the counting process.

**Lemma C.3.** *We have for all  $t, s \geq 0$  such that  $s \geq t$*

$$\begin{aligned} \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] &= N_t^a - N_t^b + \int_0^t \int_{t-r}^{s-r} \Xi(u) du (dN_r^a - dN_r^b) \\ &\quad - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{a/a}(x) dx du dN_r^a - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{b/a}(x) dx du dN_r^a \\ &\quad + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{a/b}(x) dx du dN_r^b + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x) \Phi^{b/b}(x) dx du dN_r^b, \end{aligned}$$

*Proof.* Applying Lemma C.1 along with Lemma C.2, the result is straightforward. □

The next lemma derives the limit of the conditional expectation of the process  $N_s^a - N_s^b$ .

**Lemma C.4.** *For all  $t \geq 0$ , we have*

$$\begin{aligned} \lim_{s \rightarrow \infty} \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] &= N_t^a - N_t^b + \int_0^t \left( \int_{t-x}^{\infty} \Xi(u) du \right) (dN_x^a - dN_x^b) \\ &\quad - \int_0^t \left( \int_{t-x}^{\infty} \int_0^{t-x} \Xi(u-v) (\Phi^{a/a} + \Phi^{b/a})(v) dv du \right) dN_x^a \\ &\quad + \int_0^t \left( \int_{t-x}^{\infty} \int_0^{t-x} \Xi(u-v) (\Phi^{a/b} + \Phi^{b/b})(v) dv du \right) dN_x^b, \end{aligned}$$

where  $\Xi := \Psi^{a/a} - \Psi^{b/a} = \Psi^{b/b} - \Psi^{a/b}$ .

*Proof.* Taking the limit as  $s$  tends to infinity of Lemma C.2 we obtain the almost sure convergence of  $\int_t^s \mathbb{E}[\lambda_u^a -$

$\lambda_u^b | \mathcal{F}_t] du$  towards

$$\begin{aligned} \int_t^\infty \mathbb{E}[\lambda_u^a - \lambda_u^b | \mathcal{F}_t] du &= \int_0^t \left( \int_{t-x}^\infty \Xi(u) du \right) (dN_x^a - dN_x^b) \\ &\quad - \int_0^t \left( \int_{t-x}^\infty \int_0^{t-x} \Xi(u-v)(\Phi^{a/a} + \Phi^{b/a})(v) dv du \right) dN_x^a \\ &\quad + \int_0^t \left( \int_{t-x}^\infty \int_0^{t-x} \Xi(u-v)(\Phi^{a/b} + \Phi^{b/b})(v) dv du \right) dN_x^b, \end{aligned}$$

where we have used that  $\Xi$  is integrable. Finally, using Lemma C.1 with the above yields the result.  $\square$

In order to show that the price is a martingale, we establish some elementary results about the convergence of the conditional expectations.

**Lemma C.5.** *For all  $t, s \geq 0$  such that  $s \geq t$ , we write  $Z_s^t := \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t]$  and  $Z^t := \lim_{s \rightarrow \infty} \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t]$ . Then, we have that*

- (i) *the random variable  $Z^t$  is integrable,*
- (ii) *the random variables  $(Z_s^t)$  converge almost surely to  $Z^t$ ,*
- (iii) *there exists an integrable random variable  $Y^t$  such that for all  $s \geq t$ ,  $|Z_s^t| \leq Y^t$ .*

*Proof.* By Lemma C.3, we have for all  $t, s \geq 0$  such that  $s \geq t$

$$\begin{aligned} Z_s^t = \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] &= N_t^a - N_t^b + \int_0^t \int_{t-r}^{s-r} \Xi(u) du (dN_r^a - dN_r^b) \\ &\quad - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x)\Phi^{a/a}(x) dx du dN_r^a - \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x)\Phi^{b/a}(x) dx du dN_r^a \\ &\quad + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x)\Phi^{a/b}(x) dx du dN_r^b + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} \Xi(u-x)\Phi^{b/b}(x) dx du dN_r^b. \end{aligned}$$

Therefore, writing  $|A|$  for the matrix of absolute values of all entries of  $A$ , we have

$$\begin{aligned} |Z_s^t| &\leq N_t^a + N_t^b + \int_0^t \int_{t-r}^{s-r} |\Xi(u)| du (dN_r^a + dN_r^b) \\ &\quad + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} |\Xi(u-x)\Phi^{a/a}(x)| dx du dN_r^a + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} |\Xi(u-x)\Phi^{b/a}(x)| dx du dN_r^a \\ &\quad + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} |\Xi(u-x)\Phi^{a/b}(x)| dx du dN_r^b + \int_0^t \int_{t-r}^{s-r} \int_0^{t-r} |\Xi(u-x)\Phi^{b/b}(x)| dx du dN_r^b \\ &\leq N_t^a + N_t^b + \int_0^t \int_{t-r}^\infty |\Xi(u)| du (dN_r^a + dN_r^b) \\ &\quad + \int_0^t \int_{t-r}^\infty \int_0^{t-r} |\Xi(u-x)\Phi^{a/a}(x)| dx du dN_r^a + \int_0^t \int_{t-r}^\infty \int_0^{t-r} |\Xi(u-x)\Phi^{b/a}(x)| dx du dN_r^a \\ &\quad + \int_0^t \int_{t-r}^\infty \int_0^{t-r} |\Xi(u-x)\Phi^{a/b}(x)| dx du dN_r^b + \int_0^t \int_{t-r}^\infty \int_0^{t-r} |\Xi(u-x)\Phi^{b/b}(x)| dx du dN_r^b. \end{aligned}$$

Writing  $Y^t$  for the last random variable,  $Y^t$  is independent from  $s$ , positive and we can easily check that it is of finite expectation since  $\Xi$  is integrable. This proves (iii). Furthermore, we also have  $|Z^t| \leq Y^t$  so that  $Z^t$  is integrable and (i) is verified. Finally, Lemma C.4 proved point (ii).  $\square$

**Lemma C.6.** *The price is a martingale.*

*Proof.* Since, for all  $t, \tau \geq 0$  such that  $\tau \leq t$  we have

$$\begin{aligned}\mathbb{E}[P_t | \mathcal{F}_\tau] &= P_0 + \Lambda \text{diag}(\nu) \mathbb{E}[\lim_{s \rightarrow \infty} \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] | \mathcal{F}_\tau] \\ &= P_0 + \Lambda \text{diag}(\nu) \lim_{s \rightarrow \infty} \mathbb{E}[\mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t] | \mathcal{F}_\tau] \\ &= P_0 + \Lambda \text{diag}(\nu) \lim_{s \rightarrow \infty} \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_\tau] \\ &= P_\tau,\end{aligned}$$

where the second equality is an application of the dominated convergence theorem with the random variables  $Z_s^t := \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t]$  and  $Z^t := \lim_{s \rightarrow \infty} \mathbb{E}[N_s^a - N_s^b | \mathcal{F}_t]$ , the conditions being met by Lemma C.5.  $\square$

We can simplify the expression of the price process using martingale-admissibility. This is the topic of the next lemma.

**Lemma C.7.** *The price process is of the form*

$$P_t = P_0 + \Lambda \text{diag}(\nu) \int_0^t \zeta(t-s)(dN_s^a - dN_s^b),$$

where  $\zeta(t) = \mathbb{I}_d + \int_{t-x}^\infty \Xi(u) du + \int_{t-v}^\infty \int_0^{t-v} \Xi(u-v)(\Phi^{a/a} + \Phi^{b/a})(v) dv du$ . Furthermore,  $K$  is almost everywhere differentiable.

*Proof.* By assumption, we know there exists some cross-impact kernel  $K$  such that for all  $t \geq 0$

$$P_t = P_0 + \int_0^t K(t-s)(dN_s^a - dN_s^b).$$

Therefore Lemma C.4 implies that for all  $t \geq 0$

$$\int_{t-v}^\infty \int_0^{t-v} \Xi(u-v)(\Phi^{a/a} + \Phi^{b/a})(v) dv du = \int_{t-v}^\infty \int_0^{t-v} \Xi(u-v)(\Phi^{a/b} + \Phi^{b/b})(v) dv du.$$

And Lemma C.4 yields

$$P_t = P_0 + \Lambda \text{diag}(\nu) \int_0^t \zeta(t-s)(dN_s^a - dN_s^b),$$

with  $\zeta(t) = \mathbb{I}_d + \int_{t-x}^\infty \Xi(u) du + \int_{t-v}^\infty \int_0^{t-v} \Xi(u-v)(\Phi^{a/a} + \Phi^{b/a})(v) dv du$ . In particular, we must have  $\Lambda \text{diag}(\nu) \zeta = K$  almost everywhere and  $K$  is almost everywhere differentiable.  $\square$

To obtain a simpler expression for  $K$  and the other results of the proposition, we adapt the proof of Proposition 3.3 of [14] to our structure of Hawkes processes. At time  $t \geq 0$ , either:

- there are no market orders and the price is differentiable with  $P'_t = \int_0^t K'(t-s)d(N^a - N^b)_s$ ;
- there is a buy market order on some asset, say Asset  $i$ , which happens with intensity  $\lambda_t^{b,i}$  and yields a price jump of  $K(0)\nu_i e_i$ ;
- there is a sell market order on some asset, say Asset  $i$ , which happens with intensity  $\lambda_t^{b,i}$  and yields a price jump of  $-K(0)\nu_i e_i$ .

Therefore, we have for all  $t \geq 0$

$$\begin{aligned}
\lim_{h \rightarrow 0} \frac{\mathbb{E}[P_{t+h} | \mathcal{F}_t] - P_t}{h} &= \int_0^t K'(t-s)(dN_s^a - dN_s^b) + \sum_{i=1}^d K(0) \nu_i e_i (\lambda_t^{a,i} - \lambda_t^{b,i}) \\
&= \int_0^t K'(t-s)(dN_s^a - dN_s^b) + K(0) \text{diag}(\nu) (\lambda_t^a - \lambda_t^b) \\
&= \int_0^t K'(t-s)(dN_s^a - dN_s^b) - K(0) \text{diag}(\nu) \int_0^t (\Phi^{b/b} - \Phi^{a/b})(t-s) dN_s^b \\
&\quad + K(0) \text{diag}(\nu) \int_0^t (\Phi^{a/a} - \Phi^{b/a})(t-s) dN_s^a \\
&= 0,
\end{aligned}$$

where the last equality holds since the price is a martingale by Lemma C.6. Thus, we must have for all  $t \geq 0$

$$\begin{aligned}
K'(t) &= -K(0) \text{diag}(\nu) (\Phi^{b/b}(t) - \Phi^{a/b}(t)) \\
K'(t) &= -K(0) \text{diag}(\nu) (\Phi^{a/a}(t) - \Phi^{b/a}(t)).
\end{aligned}$$

To conclude, we need to derive the expression of  $K(0)$ . Since  $\Phi^{b/b} - \Phi^{a/a}$  is integrable,  $K$  converges to a finite limit and by definition of the permanent cross-impact matrix:  $\lim_{t \rightarrow \infty} K(t) = \Lambda$ . Thus

$$\lim_{t \rightarrow \infty} K(t) = K(0) \text{diag}(\nu) \left( \mathbb{I}_d - \int_0^\infty (\Phi^{b/b}(s) - \Phi^{a/b}(s)) ds \right) = \Lambda.$$

Since by Assumption 1 we have  $\rho(\int_0^\infty \Phi(s) ds) < 1$ , we get

$$K(0) = \Lambda \left( \mathbb{I}_d - \int_0^\infty (\Phi^{b/b}(s) - \Phi^{a/b}(s)) ds \right)^{-1} \text{diag}(\nu)^{-1}.$$

Since the matrix  $\Lambda$  is non-singular by Assumption 3 so is the matrix  $K(0)$  and we must also have

$$\Phi^{b/b} - \Phi^{a/b} = \Phi^{a/a} - \Phi^{b/a}.$$

This completes the proof of Proposition 2.1.

## C.2 Proof of Proposition 2.2

Because of the autocovariance structure of Hawkes processes, this section makes extensive use of Fourier transforms of measures. We recall from the notation section that the Fourier transform of a measure  $\kappa$  at  $\omega \in \mathbb{C}$  is defined, when the integral converges, as

$$\hat{\kappa}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \kappa(dt).$$

Writing  $\theta := (\mathbb{I}_d - \Phi^{a/a})\mu + \Phi^{a/b}\mu$  for the average of the intensity of the stationary version of the Hawkes process we know from [3] that

$$\hat{\Omega}^N(\omega) = (\mathbb{I}_d - \hat{\Phi}(\omega))^{-1} \text{diag}(\theta) ((\mathbb{I}_d - \hat{\Phi}(-\omega))^{-1})^\top.$$

In particular, the Fourier transform of the reduced covariance measure  $\Omega$  has integrable entries.

We begin the proof of Proposition 2.2 with a useful lemma.

**Lemma C.8.** *For all  $\omega \in \mathbb{R}$ , we have*

$$(\widehat{K}'(\omega) + K(0))\widehat{\Omega}(\omega)(\widehat{K}'(\omega) + K(0))^* = \Lambda\Omega_\infty\Lambda^\top.$$

*Proof.* Writing  $\theta := (\mathbb{I}_d - \Phi^{a/a})\mu + \Phi^{a/b}\mu$  for the average of the intensity of the stationary version of the Hawkes process we know from [3] that

$$\widehat{\Omega}^{\tilde{N}}(\omega) = (\mathbb{I} - \widehat{\Phi}(\omega))^{-1} \text{diag}(\theta)((\mathbb{I} - \widehat{\Phi}(-\omega))^{-1})^\top.$$

Therefore we have

$$\begin{aligned} \widehat{\Omega}^{\tilde{N}}(\omega) &= (\mathbb{I}_d - \widehat{\Phi}(\omega))^{-1} \text{diag}(\theta)((\mathbb{I}_d - \widehat{\Phi}(\omega))^{-1})^* \\ \widehat{\Omega}(\omega) &= (\mathbb{I}_d, -\mathbb{I}_d)\widehat{\Omega}^{\tilde{N}}(\omega)(\mathbb{I}_d, -\mathbb{I}_d)^\top. \end{aligned}$$

Using pseudo-inverses we obtain

$$\begin{aligned} \widehat{\Omega}(\omega)^{-1} &= \frac{1}{4}(\mathbb{I}_d, -\mathbb{I}_d)(\widehat{\Omega}^{\tilde{N}}(\omega))^{-1}(\mathbb{I}_d, -\mathbb{I}_d)^\top \\ &= (\mathbb{I}_d - \widehat{\varphi}(\omega))^* \text{diag}(\theta)^{-1}(\mathbb{I}_d - \widehat{\varphi}(\omega)) \\ &= (\mathbb{I}_d - \widehat{\varphi}(\omega))^*((\mathbb{I}_d - \widehat{\varphi}(0))^{-1})^*\Omega_\infty^{-1}(\mathbb{I}_d - \widehat{\varphi}(\omega))^{-1}(\mathbb{I}_d - \widehat{\varphi}(0)), \end{aligned}$$

where we have introduced  $\Omega_\infty := \widehat{\Omega}(0)$ . Therefore, since  $\Lambda$  is real, by Proposition 2.1 we get

$$(\widehat{K}'(\omega) + K(0))\widehat{\Omega}(\omega)(\widehat{K}'(\omega) + K(0))^* = \Lambda\Omega_\infty\Lambda^\top.$$

□

We now prove Proposition 2.2. First, note that  $\widehat{\Omega}$  admits a spectral factor. since it is positive-definite almost everywhere on the unit circle. Furthermore, by Theorem B.1 and Assumption 4,  $\varphi$  is square-integrable and causal. Thus, Lemma C.8 shows that  $(\mathbb{I}_d - \widehat{\varphi})^{-1} \text{diag}(\sqrt{\theta})$  is a spectral factor of  $\widehat{\Omega}$ . Since there exists a spectral factor, the Paley-Wiener condition Equation (8) is necessarily satisfied.

As  $\widehat{\Omega}$  satisfies the conditions of Theorem B.2, let  $\mathcal{L}$  be any spectral factor of  $\widehat{\Omega}$ . Then, since the products of two functions of causal inverse Fourier transform also has a causal inverse Fourier transform, the inverse Fourier transform of  $\widehat{K}'\mathcal{L}$  is causal. Furthermore, since  $\widehat{K}'$  and  $\mathcal{L}$  are both integrable functions, their product is square-integrable. Thus, the product satisfies the first condition of Theorem B.1 and therefore the product belongs to the Hardy space  $\mathbb{H}^2$ . Therefore, by Lemma C.8,  $\widehat{K}'\mathcal{L}$  is a spectral factor of the positive-definite matrix  $\Lambda\Omega_\infty\Lambda^\top$ . It follows by the uniqueness of spectral factors up to a unitary matrix that there exists some unitary matrix  $\mathcal{O}$  such that for almost all  $\omega \in \mathbb{R}$

$$\widehat{K}'(\omega) = \mathcal{G}\mathcal{O}\mathcal{L}(\omega)^{-1} - K(0),$$

where  $\mathcal{G}\mathcal{G}^\top = \Lambda\Omega_\infty\Lambda^\top$ . This concludes the proof.

### C.3 Proof of Proposition 2.3

The immediate cross-impact matrix  $K(0)$  describes how trades push prices on very short time scales. As such, we intuitively expect that it must be constrained to prevent pair-trading arbitrage. In fact, we show in this section that it can be completely characterised. The first lemma shows that the immediate cross-impact matrix must be symmetric non-negative.

**Lemma C.9.** *The immediate cross-impact matrix  $K(0)$  is symmetric and non-negative.*

*Proof.* This proof uses no-arbitrage arguments. It is inspired from [20] but adapted here since assumptions are slightly different. We consider a deterministic trading strategy ending at time  $T$  given by the function  $f: [0, T] \rightarrow \mathbb{R}^d$  which determines the buy and sell market orders according to Definition 1.2. We re-write the cross-impact contributions as  $K(t) = M + H(t)$ , where  $M$  is the immediate impact matrix and  $H(0) = 0$ . From Proposition 2.1 we know that  $H$  is continuous at zero. With these conventions, the average cost  $C(f)$  of the trading strategy  $f$  given in Definition 1.2 is written as

$$C(f) = \int_0^T f(t)^\top \int_0^t Mf(s) ds dt + \int_0^T f(t)^\top \int_0^t H(t-s)f(s) ds dt =: C_i(f) + C_s(f),$$

where we have split costs into two parts. The first,  $C_i(f)$ , represents immediate impact costs and the second,  $C_s(f)$ , represents the rest. Then

$$C_i(f) = \frac{1}{2} \sum_{i=1}^d M_{ii} \|f_i\|_1^2 + \sum_{i \neq j}^d M_{ij} \int_0^T \int_0^t f_i(t) f_j(s) ds dt.$$

Similarly, for the other impact costs, we have

$$C_s(f) = \sum_{i=1}^d \int_0^T \int_0^t f_i(t) H_{ii}(t-s) f_i(s) ds dt + \sum_{i \neq j}^d \int_0^T \int_0^t f_i(t) H_{ij}(t-s) f_j(s) ds dt.$$

We choose two distinct assets, Asset  $a$  and Asset  $b$  and consider, as in [20], a round-trip pair-trading strategy of the following form, where  $v_p, v_q \in \mathbb{R}$ :

$$f_p(t) := \begin{cases} v_p & \text{for } 0 \leq t \leq T/3 \\ 0 & \text{for } T/3 \leq t \leq 2T/3, \\ -v_p & \text{for } 2T/3 \leq t \leq T \end{cases} \quad f_q(t) := \begin{cases} v_q & \text{for } 0 \leq t \leq T/3 \\ -v_q & \text{for } T/3 \leq t \leq 2T/3 \\ 0 & \text{for } 2T/3 \leq t \leq T \end{cases}.$$

This strategy only trades Asset  $p$  and Asset  $q$ , so that for all other Asset  $i$ ,  $f_i = 0$ . This is a round-trip strategy since  $\int_0^T f = 0$ . Then, the immediate impact costs contribution is

$$C_i(f) = \frac{T^2}{18} (M_{pq} - M_{qp}) v_p v_q.$$

For the other impact costs, we have

$$C_s(f) = \int_0^T \int_0^t f_a(t) H_{pp}(t-s) f_p(s) ds dt + \int_0^T \int_0^t f_q(t) H_{qq}(t-s) f_q(s) ds dt \\ + \int_0^T \int_0^t f_p(t) H_{pq}(t-s) f_q(s) ds dt + \int_0^T \int_0^t f_q(t) H_{qp}(t-s) f_q(s) ds dt.$$

Therefore, since  $H(0) = 0$  and  $H$  is continuous at  $t = 0$ , for small enough execution times  $T$  we have that for all  $t \in [0, T]$ ,  $|H_{ij}(t)| \leq \epsilon$  for all  $(i, j) \in \{p, q\}^2$ . Thus, combining both impact terms, we get

$$\frac{C(f)}{T^2} \leq \frac{\nu_a \nu_b}{18} (M_{qp} - M_{pq}) + \epsilon.$$

So, unless  $M_{pq} = M_{qp}$ , the volumes  $\nu_p$  and  $\nu_q$  can be chosen so that trading costs of this round trip strategy are negative. Thus, if  $M_{pq} \neq M_{qp}$ , arbitrages are possible. Therefore the immediate cross-impact matrix  $M = K(0)$  is necessarily symmetric.

Since  $K(0)$  is symmetric and  $K$  is continuous by Proposition 2.1, Lemma 2.8 of [1] implies that  $K(0)$  is non-negative definite and hence non-negative. This completes the proof.  $\square$

The previous results show that the immediate cross-impact matrix is symmetric non-negative. By further using the price dynamics, we are able to relate it to the instantaneous covariance matrix of prices and order flows.

We have

$$dP_t = K(0) \text{diag}(\nu)(dN_t^b - dN_t^a).$$

Taking the predictable quadratic variation of the processes, we obtain

$$d\langle P, P \rangle_t = K(0) \text{diag}(\nu^2) \text{diag}(\lambda_t^a + \lambda_t^b) K(0)^\top dt$$

Taking expectations on both sides and writing the return covariance matrix  $\Sigma_t dt := \mathbb{E}[d\langle P, P \rangle_t]$  and the average intensity  $\theta_t := \mathbb{E}[\lambda_t^b] = \mathbb{E}[\lambda_t^a]$  we get

$$\frac{1}{2} \Sigma_t = K(0) \text{diag}(\nu^2 \theta_t) K(0)^\top.$$

Since this holds for all  $t$ , and  $\theta_t \xrightarrow[t \rightarrow \infty]{} \theta = (\mathbb{I}_d - \Phi^{a/a})\mu + \Phi^{a/b}\mu$ , it must hold as  $t$  tends to infinity and we get

$$\frac{1}{2} \Sigma = K(0) \text{diag}(\theta \nu^2) K(0)^\top,$$

where  $\Sigma = \lim_{t \rightarrow \infty} \mathbb{E}[d\langle P_t, P_t \rangle]$ , which is well-defined by passing to the limit in the above. Since, by Lemma C.9,  $K(0)$  is symmetric non-negative and satisfies the above, it must be of the form (see for example Proposition 3 of [22])

$$K(0) = \frac{1}{\sqrt{2}} (\mathcal{L}_0^{-1})^\top \sqrt{\mathcal{L}_0^\top \Sigma \mathcal{L}_0} \mathcal{L}_0^{-1},$$

where  $\mathcal{L}$  is given in the proposition. This completes the proof of Proposition 2.3.

#### C.4 Proof of Proposition 2.4

As in the proof of Appendix C.3, we consider a trading strategy ending at time  $T$  given by the function  $f: [0, T] \rightarrow \mathbb{R}^d$ . We re-write the cross-impact contributions as  $K(t) = \Lambda + \Gamma(t)$ , where  $\Gamma(t) \xrightarrow[t \rightarrow \infty]{} 0$ . Then, average cost  $C(f)$  of the trading strategy  $f$  given by Definition 1.2 can be re-written as

$$C(f) := \int_0^T f(t)^\top \int_0^t \Lambda f(s) ds dt + \int_0^T f(t)^\top \int_0^t \Gamma(t-s) f(s) ds dt = C_p(f) + C_t(f),$$

where we have split into permanent and temporary impact costs. Then

$$C_p(f) = \frac{1}{2} \sum_{i=1}^d \Lambda_{ii} \|f_i\|_1^2 + \sum_{i \neq j}^d \Lambda_{ij} \int_0^T \int_0^t f_i(t) f_j(s) ds dt.$$

Similarly, for the temporary impact costs, we have

$$C_t(f) = \sum_{i=1}^d \int_0^T \int_0^t f_i(t) \Gamma_{ii}(t-s) f_i(s) ds dt + \sum_{i \neq j}^d \int_0^T \int_0^t f_i(t) \Gamma_{ij}(t-s) f_j(s) ds dt.$$

Consider two distinct assets, Asset  $a$  and Asset  $b$ . Consider, as in [20], a round-trip trading strategy of the following form, here  $v_p, v_q \in \mathbb{R}$ :

$$f_a(t) := \begin{cases} v_p & \text{for } 0 \leq t \leq T/3 \\ 0 & \text{for } T/3 \leq t \leq 2T/3, \\ -v_p & \text{for } 2T/3 \leq t \leq T \end{cases} \quad f_q(t) := \begin{cases} v_q & \text{for } 0 \leq t \leq T/3 \\ -v_q & \text{for } T/3 \leq t \leq 2T/3 \\ 0 & \text{for } 2T/3 \leq t \leq T \end{cases}.$$

Then, the permanent impact costs contribution is of the form

$$C_p(f) = (\Lambda_{pq} - \Lambda_{qp}) \frac{T^2}{18} v_p v_q.$$

Similarly, for the temporary impact costs, we have

$$\begin{aligned} C_t(f) &= \int_0^T \int_0^t f_p(t) \Gamma_{pp}(t-s) f_p(s) ds dt + \int_0^T \int_0^t f_q(t) \Gamma_{qq}(t-s) f_q(s) ds dt \\ &\quad + \int_0^T \int_0^t f_p(t) \Gamma_{pq}(t-s) f_q(s) ds dt + \int_0^T \int_0^t f_q(t) \Gamma_{qp}(t-s) f_p(s) ds dt. \end{aligned}$$

Therefore, since  $\Gamma$  is power-law with all exponents strictly below 1, we have

$$\frac{C(f)}{T^2} \underset{T \rightarrow \infty}{=} \frac{v_p v_q}{18} (\Lambda_{qp} - \Lambda_{pq}) + o(1),$$

so that if  $\Lambda_{qp} \neq \Lambda_{pq}$ , arbitrages are possible. Therefore  $\Lambda$  is necessarily symmetric.

We now show that  $\Lambda$  is non-negative. Let  $\tau \in \mathbb{R}$  and  $\eta \in \mathbb{R}^d$  and consider a trading strategy  $f$  which buys portfolio  $\eta$  and waits  $\tau$  units of time to sell it. Then, the cost of this trading strategy is

$$C(f) = \eta^\top (\Lambda + \Gamma(\tau)) \eta,$$

and by no-arbitrage,  $C(f) \geq 0$ . Therefore, for all  $\tau \in \mathbb{R}$  and  $\eta \in \mathbb{R}^d$ , we have

$$\eta^\top \Lambda \eta \geq -\eta^\top \Gamma(\tau) \eta,$$

so that using the fact that  $\Gamma(\tau) \xrightarrow{\tau \rightarrow \infty} 0$ ,  $\Lambda$  is non-negative. This concludes the proof of Proposition 2.4.

## C.5 Proof of Proposition 2.5

To prove Proposition 2.5, we proceed in three steps. First, we use polarization to show that it suffices to prove the result for  $\xi^* \mathcal{Z} \xi$  for any  $\xi \in \mathbb{C}^d$ . Second, we use sufficient conditions for smoothness on characteristic functions. Finally, we show that these conditions are satisfied in our setting.

We begin with a polarization identity. Since  $\mathcal{Z}$  is a continuous positive definite function, Theorem 2.10 of [1] shows that for every  $\xi \in \mathbb{C}^d$  the continuous function  $\mathcal{Z}_\xi: t \mapsto \xi^* \mathcal{Z}(t) \xi$  is positive definite. Furthermore, for all  $1 \leq a, b \leq d$ , the component  $e_b^\top \mathcal{Z}(t) e_a$  is equal to

$$\frac{1}{2} (\mathcal{Z}_{e_a+e_b}(t) - i \mathcal{Z}_{e_a-i e_b}(t) - (1-i) \mathcal{Z}_{e_a}(t) - (1-i) \mathcal{Z}_{e_b}(t)).$$

Thus it suffices to prove the result for  $\mathcal{Z}_\xi$  for  $\xi \in \mathbb{C}^d$ , which we show below.

For all  $\xi \in \mathbb{C}^d$ , we introduce the measure  $\mathbb{Q}_\xi := \frac{\xi^* \mathbb{M} \xi}{\xi^* \mathbb{M}(\mathbb{R}) \xi}$ . For any  $\xi \in \mathbb{C}^d$ ,  $\mathbb{Q}_\xi$  is indeed a probability measure since the matrix-valued measure  $\mathbb{M}$  is non-negative definite and of finite total variation. Thus, the function  $\mathcal{Z}_\xi$  is the characteristic function associated to the probability measure  $\mathbb{Q}_\xi = \frac{\xi^* \mathbb{M} \xi}{\xi^* \mathbb{M}(\mathbb{R}) \xi}$ . We now use Theorem 2.3.1 of [17] which gives a sufficient condition for the smoothness of a characteristic function. To apply the theorem, we must show that for all  $\xi \in \mathbb{C}^d$  we have

$$\liminf_{t \rightarrow 0} \frac{|\mathcal{Z}_\xi(2t) - \mathcal{Z}_\xi(0) + \mathcal{Z}_\xi(-2t) - \mathcal{Z}_\xi(0)|}{4t^2} < \infty.$$

Using the definition of  $\mathcal{Z}_\xi$ , the above condition is equivalent to

$$\liminf_{t \rightarrow 0} \frac{|\xi^* (K(2t) - K(0) + K(2t)^\top - K(0)^\top) \xi|}{4t^2} < \infty.$$

However, by Assumption 4,  $K$  and its derivative are continuously differentiable at zero and Proposition 2.3 shows that  $K(0)$  is symmetric, so that

$$\frac{K(2t) - K(0) + K(2t)^\top - K(0)^\top}{4t^2} \xrightarrow[t \rightarrow 0]{} \frac{1}{2t} (K'(0) + K'(0)^\top) + K''(0) + K''(0)^\top + o(1).$$

Therefore the condition will be satisfied and the proposition proven if  $K'(0) = -K'(0)^\top$ , that is if  $K'(0)$  is antisymmetric. Then Theorem 2.3.1 of [17] yields that  $\mathcal{Z}_\xi$  is twice differentiable and the integrals converge absolutely. To prove this result, we will use the no-arbitrage condition and the smoothness of the cross-impact kernel around zero.

As in the proofs of Appendices C.3 and C.4, we consider a trading strategy ending at time  $T$  given by the

function  $f: [0, T] \rightarrow \mathbb{R}^d$ . We re-write the cross-impact contributions as  $K(t) = K(0) + tK'(0) + \mathcal{R}(t)$ , where  $\mathcal{R}$  is such that  $\mathcal{R}(0) = 0$ ,  $\mathcal{R}$  is continuously differentiable at zero and  $\mathcal{R}(t) \underset{t \rightarrow 0}{=} o(t)$ . For convenience, we write  $M(t) := K(0) + tK(0)$ . Then the trading cost of the trading strategy is

$$C(f) := \int_0^T f(t)^\top \int_0^t M(t-s)f(s)ds dt + \int_0^T f(t)^\top \int_0^t \mathcal{R}(t-s)f(s)ds dt = C_1(f) + C_2(f).$$

Where we have split contributions. Consider two distinct assets, Asset  $a$  and Asset  $b$  and the round-trip trading strategy of the following form, where  $v_a, v_b \in \mathbb{R}$ :

$$f_a(t) := \begin{cases} v_a & \text{for } 0 \leq t \leq T/3 \\ 0 & \text{for } T/3 \leq t \leq 2T/3, \\ -v_a & \text{for } 2T/3 \leq t \leq T \end{cases} \quad f_b(t) := \begin{cases} v_b & \text{for } 0 \leq t \leq T/3 \\ -v_b & \text{for } T/3 \leq t \leq 2T/3 \\ 0 & \text{for } 2T/3 \leq t \leq T \end{cases}.$$

This strategy only trades Asset  $a$  and Asset  $b$ , i.e. for all other Asset  $i$ ,  $f_i = 0$ . This is a round-trip strategy since  $\int_0^T f = 0$ . Then, computing contributions as in Lemma C.9 and using the fact that  $K(0)$  is symmetric, we obtain

$$C_1(f) = \frac{-5T^3}{162} v_a v_b (K'(0)_{ab} + K'(0)_{ba}).$$

For the other contribution, we have

$$\begin{aligned} C_2(f) &= \int_0^T \int_0^t f_a(t) \mathcal{R}_{aa}(t-s) f_a(s) ds dt + \int_0^T \int_0^t f_a(t) \mathcal{R}_{bb}(t-s) f_a(s) ds dt \\ &\quad + \int_0^T \int_0^t f_a(t) \mathcal{R}_{ab}(t-s) f_b(s) ds dt + \int_0^T \int_0^t f_b(t) \mathcal{R}_{ba}(t-s) f_a(s) ds dt. \end{aligned}$$

Therefore, since  $\mathcal{R}(0) = 0$  and  $\mathcal{R}(t) \underset{t \rightarrow 0}{=} o(t)$ , for small enough execution times  $T$  we have that for all  $t \in [0, T]$ ,  $|\mathcal{R}_{ij}(t)| \leq \epsilon t$  for all  $(i, j) \in \{a, b\}^2$ . Thus, we then have, where  $c_1 > 0$  is a constant independent of  $f$ ,  $T$  and  $\epsilon$

$$C_2(f) \leq \epsilon c_1 T^3.$$

Thus, combining both terms we obtain

$$\frac{C(f)}{T^3} \leq c_2 v_a v_b (K'(0)_{ab} + K'(0)_{ba}) + \epsilon,$$

where  $c_2 \neq 0$  is a constant independent of  $f$ ,  $T$  and  $\epsilon$ . Therefore, unless  $K'(0)_{ab} = -K'(0)_{ba}$ , the round-trip strategy yields negative costs and arbitrages are possible. Therefore  $K'(0)$  is necessarily antisymmetric. The proposition then follows.

## D Calibration methodology details

This section details the calibration methodology. The objective is to calibrate the cross-impact kernels  $K^1$  and  $K^2$  on empirical data. Appendix D.1 gives additional details on the dataset. Appendix D.2 details the methodology for the estimation of the empirical observables  $\Sigma$  and  $\Omega$ . Appendix D.3 explains how spectral factors of the reduced covariance measure are computed. Finally, Appendix D.4 explains the construction of

both kernels.

## D.1 Data preparation and processing

The data is processed according to the procedure outlined in [22]. For the reader's convenience, we recall some key elements here. The two instruments selected are the leading and third month E-Mini Futures. Prices and trades are gathered from anonymous trades and quotes data from the CME and prices are taken as the mid-price of the best bid and ask prices of each instrument. To avoid stationarity issues, we remove data outside the commonly traded hours of both instruments and the first and last 30 minutes of the trading period. Data ranges from January 2015 to December 2018.

## D.2 Estimation of empirical observables

Using the data outlined in Appendix D.1, we detail here the estimation methodology for  $\Sigma$  and  $\Omega$ . To simplify computations and reduce noise, we aggregate order flows and prices by bins of 1 seconds. Each bin contains the opening and closing price  $p_t^o$  and  $p_t^c$  as well as the total signed order flow  $q_t$ . Both of these random variables have zero mean. Using these conventions, we use naive statistical estimators to compute the price-covariance and the order flow auto-covariance, namely, for a daily timeseries of prices and signed order flows  $\{p_t\}_{t=1}^T$  and  $\{q_t\}_{t=1}^T$  at bins of one second:

$$\begin{aligned}\Sigma &= \frac{1}{T-1} \sum_{t=1}^T (p_t^c - p_t^o)(p_t^c - p_t^o)^\top \\ \Omega(\tau) &= \frac{1}{T} \sum_{t=1}^T q_{t+\tau} q_t^\top.\end{aligned}$$

This procedure is averaged across days to obtain an empirical estimate of  $\Sigma$  and  $\Omega$  using approximately 800 days of available data. For stationary ergodic point processes, these averages do converge towards the theoretical price-covariance and reduced covariance measure by Proposition 8.3.1 of [8].

## D.3 Estimation of a spectral factor of $\Omega$

Once  $\Sigma$  and  $\Omega$  have been estimated, the cross-impact kernel  $K^1$  can be computed by solving numerically Equation (5) and setting the boundary conditions imposed by arbitrage-admissibility. To do so, we need to compute a spectral factor of  $\Omega$ .

To solve Equation (9), we use the SBR2 algorithm [18, 31], a polynomial eigenvalue decomposition method, to compute a numerical approximation of a spectral factor of  $\Omega$ ,  $\mathcal{L}$ . The implementation details follow the simple version of this algorithm (and not its subsequent improvements) outlined in Equation (9). The algorithm was tested on the examples provided in both papers and results were similar to those reported in the papers. To estimate the accuracy of the spectral decomposition, we computed the Frobenius norm of the error matrix  $\frac{\|\Omega(z) - \mathcal{L}(z)\mathcal{L}(1/z)^*\|_F}{\|\Omega(z)\|_F} \approx 6 \cdot 10^{-8}$ . Thus, numerically,  $\mathcal{L}$  is a good approximation for a spectral factor of  $\Omega$ .

A key property is that this spectral factor has the convenient form  $\mathcal{L}(z) = D(z)H(z)$  where  $H$  is a para-unitary

polynomial matrix and  $D$  is (close to) a diagonal polynomial matrix. Both  $D$  and  $H$  are outputs of the SBR2 algorithm. To check how close  $D$  is to a diagonal matrix, we compute the Frobenius norm of its off-diagonal elements relative to its diagonal elements. The numerical results yield  $10^{-15}$ , which shows that the decomposition was successful. Once this spectral factor has been obtained, we need to compute the matrix polynomial  $\mathcal{L}^{-1}$  to obtain the martingale-admissible kernel  $K^1$ .

#### D.4 Computation of the cross-impact kernels

Because of the structure of  $\mathcal{L}(z)$ , computing its inverse is straightforward. One the one hand, the inverse of the para-unitary polynomial matrix  $H(z)$  is  $H(1/z)^*$ . On the other hand, the inverse of  $D(z)$  is obtained by taking the inverse of its diagonal elements, which is straightforward using pole decomposition. This allows us to compute numerically the polynomial matrix  $\mathcal{L}^{-1}$  and, using Equation (5), the martingale-admissible kernel  $K^1$ .

The numerical computation of the poles of polynomials is done using the `residuez` function of the `signal` library of the `scipy` module [27] of the python programming language [26]. All poles found were strictly inside the unit circle. However, they were numerically close to modulus one, which is consistent with the long-range auto-correlations of  $\Omega$  reported in Figure 1.

Once the martingale-admissible kernel  $K^1$  has been obtained, we derive the nsa-admissible kernel  $K^2$  in the following manner. First, we compute the (numerical) Fourier transform of  $K^1$ . The symmetric part of this Fourier transform is then modified so that all its eigenvalues are non-negative. Finally, this clipped Fourier transform is added to the asymmetric part of the Fourier transform of  $K^1$ . The (numerical) inverse Fourier transform then yields the nsa-admissible kernel  $K^2$ .