KYBERNETIKA — VOLUME *48* (2012), NUMBER 1, PAGES 50–82

PROBABILISTIC PROPERTIES OF THE CONTINUOUS DOUBLE AUCTION

Martin Šmíd

In this paper we formulate a general model of the continuous double auction. We (recursively) describe the distribution of the model. As a useful by-product, we give a (recursive) analytic description of the distribution of the process of the best quotes (bid and ask).

Keywords: continuous double auction, limit order market, distribution

Classification: 91G80

1. INTRODUCTION

Currently, the continuous double auction (c.d.a.) is one of the most common trading mechanisms used at financial markets.

Trading according to the c.d.a. is asynchronous, i. e., agents may take their actions at any instant during the trading hours. Three types of actions may be performed:

- putting a *limit order*, i. e., an offer to buy (or sell) a certain amount of the traded commodity for no more (no less) than a certain *limit price*
- putting a *market order*, i. e., the instruction to buy (sell) a certain amount of the commodity for the best available price(s)
- cancelling a (previously submitted) pending (i. e., still unexecuted) limit order.

The collection of all pending buy (sell) limit orders is called a *buy (sell) limit order book*, the best available buying (selling) price is referred to as the *best bid* (*best ask*).

In the most recent decade, several papers on modelling the continuous double auction have appeared: from a number of similar models, let us mention the simple one by [5] with Poisson orders' arrivals and the bounded uniform relative limit prices, the static model by [4] with non-uniform continuous absolute limit prices, a discrete-price uniform model by [11] and, finally, the recent generalisation of the last model by [6] incorporating several properties of real-life limit order markets.

Even if the statistical properties of the models have been extensively researched by means of simulations and approximations (see, e. g., [11]), their rigorous probabilistic descriptions have not yet been developed.

The present paper makes a step toward this goal: a general model, covering the models by [4, 5] and [11], is formulated and a rigorous (recurrent) description of its distribution is provided.

Similarly to the covered models, unit order sizes are assumed in the general model. The evolution of the market (i. e., of the order books) is assumed to (stochastically) depend only on the history of the best quotes, which roughly corresponds to the reality in which the order books are hidden to (the majority of) the participants. The (conditional) distribution of the limit prices is assumed to be continuous; however, as demonstrated further on, the model may be easily transformed to enable discrete prices (price ticks).

Our model allows both finite and infinite numbers of the orders in the order books. Even if the assumption of infinite order books is clearly unrealistic, it brings several advantages: First, the order books never become empty given the assumption which simplifies the modelling, in cases when the probability of an empty order book is negligible anyway. Second, no arbitrary truncation point of the limit prices (as in [5]) has to be set. Third and most importantly, the infinite order books are necessary to be assumed for our model to cover the popular model by [11].[1] Moreover, the assumption of infinite books does not harm the predictions of the model much, because the probability of executing the orders lying deep in the book before they are cancelled is so small that their influence on the dynamics of the model is negligible.

The main theoretical result of the paper is the analytical description of the distribution of the market (i. e., the order books) at a given time. As a first step and simultaneously a (very useful) by-product, the distribution of the best quotes (the bid and the ask) is derived, allowing us, to some extent, to work with the model without the knowledge of the (usually inaccessible) order book data. The distribution of the order books is then described by means of the distribution of the quotes and the conditional distribution of the order books given the history of the quotes.

There are many applications of our theoretical results. Having formulas at hand, we can construct tests and/or estimators of the parameters of the model, we can design algorithms for efficient Monte Carlo simulation or we can examine further theoretical properties of the model; one of those applications may be found in the accompanying paper [9] where the tails of the price increments in the model by [11] are examined: in particular, it is proven that the tails are fat (as it was demonstrated by the simulations, see [11] and the references therein) but, quite surprisingly, the tails become thin if an initial call auction is held before the trading begins. Another theoretical application of the model could be an algorithm for optimal purchase and/or liquidation of a large amount of a stock, for instance.

This paper is organized as follows: First (Section 2), some notation and definitions of mathematical objects, uncommon in financial literature, are introduced. Further (Section 3), the general model is defined. Next (Section 4), the distribution of the model is specified. Finally (Section 5), the paper is concluded. Descriptions of the covered models by means of our notation, proofs and auxiliary theoretical results can be found in the Appendix.

[1] The reasons why [11] work with infinite order books, most likely, are the statistically-mechanical analogies used by the authors.

2. PRELIMINARIES

In the current Section, we discuss the mathematics we shall use to describe the model.

Since, in principle, there is no upper bound of the amount of pending limit orders, the mathematical objects describing the order books have to be infinitely dimensional (even in the case when we do not assume actually infinite order books). From several equivalent representations of the shapes of the order books, we have chosen (simple) point processes,[2] whose definition and basic properties are summarized in Subsection 2.1.

Since we use conditional distributions extensively in the present paper, we could not avoid using conditional independence – in Subsection 2.2, we provide a brief introduction into this notion.

Finally, we introduce some notation, specific to the present paper (Subsection 2.3).

2.1. Point Processes Basics

X is a *k-dimensional (real) point process* (p.p.) if it is a collection of random variables (points) taking values in $\mathbb{R}^k$ such that each bounded[3] set contains no more than finitely many points. If the points are distinct with probability one (w.p.1) then the p.p. is called *simple*.

Since simple point processes may be viewed as sets, we shall sometimes use set-theory notation to describe their relationships.

Moreover, since point processes are equivalent to random σ-finite atomic measures, we shall also adopt some notation from measure theory: By $|X|$ we shall denote the total number of the points of X. Under XS, where S is a Borel set, we shall understand the number of points belonging to S. By $X|_S$ we denote the restriction of a p.p. (or a measure) X to the sub-space S.

We say that a p.p. X is *finite* if $|X| < \infty$.

If there exists a value $m_i \in \mathbb{R}^M$, assigned to each point of x_i of a p.p. X defined on $\mathbb{R}^N$, $i \leq |X|$, then we say that X is a *marked point process* (m.p.p.) with marks in $\mathbb{R}^M$; in that case, the collection $(x_i)_{i \leq |X|}$ is called a *ground process of* X. Note that X may be viewed as a p.p. on $\mathbb{R}^{M+N}$ – we shall view m.p.p.'s that way in the sections that follow.

A p.p. X is a *Poisson (point) process* (P.p.p.) on $\mathbb{R}^k$ with an intensity measure μ if

(i) $XS \sim \text{Poisson}(\mu S)$ for any bounded $S \in \mathfrak{B}(\mathbb{R}^k)$ (here, $\mathfrak{B}(M)$ denotes Borel sets of M)

(ii) $XS_1, XS_2, \ldots XS_n$ are independent for any disjoint collection of Borel sets S_1, $S_2, \ldots, S_n$,

(iii) $\mu S < \infty$ for any bounded Borel set S.

It easily follows from the definition of a P.p.p. and from the additive property of the Poisson distribution that, whenever $X_1, \ldots, X_n$ are mutually independent P.p.p.'s with

[2]The point processes seem to be the most suitable for us because the theory describing them is the most developed and intuitive among other equivalent representations, e. g., stepwise functions or infinitely dimensional vectors.

[3]I.e., contained in a hypercube of a finite volume.

intensities $\mu_1, \ldots, \mu_n$, $X_1 + X_2 + \cdots + X_n$ is a P.p.p. with the intensity $\sum_i \mu_i$ (here, the sum of the point processes means the p.p. comprising of all the points of the summed processes).

2.2. Conditional Independence

The present Subsection is devoted to the notion of conditional independence which is in fact nothing else but the (ordinary) independence of conditional distributions. While the ordinary independence of random variables U_1 and U_2 may be interpreted as the non-existence of any information based on U_2 which could help us forecast U_1, the conditional independence of U_1 and U_2 given V means that all the information common to U_1 and U_2 may be explained by V. By analogy with ordinary independence, the conditional one is usually defined by means of (conditional) probabilities as follows: we say that random elements[4] $U_1, \ldots, U_n$ are conditionally independent given a random element V if

$$\mathbb{P}[U_1 \in A_1, \ldots, U_n \in A_n | V] = \prod_{i=1}^{n} \mathbb{P}[U_i \in A_i | V] \quad \text{w.p.1}$$

for any collection of measurable sets $A_1, \ldots, A_n$. In the sequel, we shall repeatedly exploit the fact that the conditional independence introduced above holds if and only if

$$\mathcal{L}\,(U_i\,|U_1, \ldots, U_{i-1}, V\,) = \mathcal{L}\,(U_i\,|V\,), \qquad 1 \leq i \leq n$$

where $\mathcal{L}\,(X\,|Y\,)$ denotes the conditional distribution of X given Y (see [3], Propositions 6.6 and 6.8).

2.3. Special Notation

At this Subsection, let us introduce the notation specific for the present work.

For any sequence $\eta_0, \eta_1, \ldots$ we put $\Delta\eta_i \stackrel{\triangle}{=} \eta_i - \eta_{i-1}$, $i \in \mathbb{N}$ (the symbol $\stackrel{\triangle}{=}$ means definition),

For any process $(x_t)_{t \geq 0}$ and each $t \in \mathbb{R}_+$, we shall write $\bar{x}_t \stackrel{\triangle}{=} (x_s)_{s \leq t}$ (i. e., its history up to t).

By writing $X = [x_1, x_2, \ldots]$ we shall say that a simple p.p. X consists of the points $x_1 < x_2 < \ldots$.

By saying that X is a P.p.p. with the density ν we shall mean that X is a P.p.p. with the intensity whose density (with respect to the Lebesgue measure) is ν.

Generalising a widely used notion, we shall say that Y is a *heterogeneous thinning (h-thinning) of a simple p.p.* $X = [x_1, x_2, \ldots]$ *with parameters* $p_1, p_2, \ldots$ if $Y \subseteq X$ and, for any $n \in \mathbb{N}$ and any distinct $i_1, i_2, \ldots i_n \in \mathbb{N}$, $i_j \leq |X|$, $j \leq n$,

$$\mathbb{P}[x_{i_1} \in Y, \ldots x_{i_n} \in Y | X] = \prod_{j=1}^{n} p_{i_j}.$$[5]

[4] By *random element* we mean any measurable mapping from an underlying probability space into a measurable space.

[5] Note that, for the definition to make sense, $p_1, p_2, \ldots$ have to be $\sigma(X)$-measurable.

3. MODEL DEFINITION

We describe the state of the market at a time $t \in \mathbb{R}_+$ by a couple

$$\Xi_t = (A_t, B_t)$$

where A_t and B_t are (finite or infinite) simple point processes with values in $\mathbb{R}$; naturally, each point of A_t or B_t stands for a limit order with a limit price equal to the location of the point.[6]

Let us agree to write $a_{t,i}$ for the ith least point of A_t and $b_{t,i}$ for the ith greatest point of B_t with the convention that $a_{t,i} = \infty$ if $|A_t| < i$ and $b_{t,i} = -\infty$ if $|B_t| < i$. Specially, let us denote

$$a_t \triangleq a_{t,1}, \qquad b_t \triangleq b_{t,1},$$

the (best) ask and (best) bid, respectively, and we naturally require that

$$b_t < a_t. \tag{1}$$

3.1. Dynamics of Ξ

To handle models with infinite order books, we cannot define the dynamics of Ξ "jump-by-jump" due to the possibly infinite number of jumps; instead, we do it (slightly less intuitively) by means of the process of the best quotes, the jumps of the quotes into the spread and the in-book order flows.

Before we formulate the definition of the dynamics of the model, let us note that the impact of an arrival of a market order is identical to that of a cancellation of the ask hence we may treat both these actions jointly, calling them *effective buy market orders*[7] – the definition of the *effective sell market* order is symmetric.[8] Further, it will be useful for us to distinguish the *in-spread limit orders* and *in-book limit orders*, the first being those whose limit prices lie between the current bid and ask, the latter being the remaining ones. Finally, let us agree to abbreviate "buy market order" as "b.m.o.", "buy limit order" as "b.l.o.", etc.

Starting with the dynamics, let the following objects be given:

- The initial value $\Xi_0 = (A_0, B_0)$, where A_0 and B_0 are simple p.p.'s on $\mathbb{R}$ fulfilling (1) for $t = 0$.

- The collection of variables
$$0 = \tau_0 < \tau_1 < \dots$$
such that $\lim_\nu \tau_\nu = \infty$ standing for jump times of the process of the best quotes.

[6] Since we allow for working with log-prices, we allow negative values of the points.

[7] Our results will not be changed if we distinguish those events; however, we handle them jointly here for notational simplicity; otherwise, two additional possible values of χ should be considered and two additional processes should be added to ξ; see the sequel for the definition of χ and ξ.

[8] We do not specifically model the *crossing limit orders*, i. e., those whose limit pricecrosses the opposite best quote (which may be submitted by mistake, for instance, see [11]) because they can be regarded as market orders (their effect is exactly the same).

- The collection of variables

$$\chi_i \in \{\mathfrak{a}^+, \mathfrak{a}^-, \mathfrak{b}^+, \mathfrak{b}^-\}, \qquad i \in \mathbb{N},$$

denoting, for each $i \in \mathbb{N}$, the type of the event happening at τ_i, where the meaning of the symbols is given according to the following list:

$\mathfrak{a}^+$ — an arrival of an effective b.m.o.

$\mathfrak{a}^-$ - an arrival of an in-spread s.l.o.

$\mathfrak{b}^+$ — an arrival of an in-spread b.l.o.

$\mathfrak{b}^-$ — an arrival of an effective s.m.o.

- The collection of variables

$$(y_i)_{i\in\mathbb{N}}, \qquad (z_i)_{i\in\mathbb{N}};$$

each y_i, z_i, respectively, $i \in \mathbb{N}$, denoting the relative limit price of an in-spread s.l.o., b.l.o., respectively (if $\chi_i \neq \mathfrak{a}^-$ then y_i equals to zero, analogously for $\mathfrak{b}^+$ and z_i).

- A collection of marked point processes describing the flows of in-book limit orders between τ_{i-1} and τ_i

$$(\mathcal{Y}_i)_{i\in\mathbb{N}}, \qquad (\mathcal{Z}_i)_{i\in\mathbb{N}}$$

defined on $(0, \Delta\tau_i) \times \mathbb{R}_+$, $(0, \Delta\tau_i) \times \mathbb{R}_-$, respectively, with marks in $\mathbb{R}_+$ where each point (θ, p, c) of $\mathcal{Y}_i$ stands for a s.l.o. submitted at the time $\tau_{i-1} + \theta$ with a limit price $a_{\tau_{i-1}} + p$ and with lifetime c, i. e., $c < \Delta\tau_i - \theta$ is interpreted as a cancellation of the order at the time $\tau_{i-1} + \theta + c$ while $c \geq \Delta\tau_i - \theta$ means that the order is still uncancelled at τ_i. Symmetrically, a point (θ, p, d) of $\mathcal{Z}_i$ means a b.l.o. submitted at the time $\tau_{i-1} + \theta$ with a limit price $b_{\tau_{i-1}} + p$ and lifetime d.

- For each $i \in \mathbb{N}$, the collections of positive variables

$$e_i = (e_{i,\nu+1})_{\nu\in\mathbb{N}}, \qquad f_i = (f_{i,\nu+1})_{\nu\in\mathbb{N}},$$

where, for each $i, \nu \in \mathbb{N}$, the variable $e_{i,\nu}$ stands for the lifetime of the ν-th least s.l.o (including the ask) present in the order book at time τ_{i-1} and where the lifetimes are measured from τ_{i-1}. Similarly, $f_{i,\nu}$ denotes the lifetime of the νth greatest b.l.o. (including the bid) present in the order book at time τ_{i-1}.

The dynamics itself is defined as follows: for any $i \in \mathbb{N}$ and any $t \in (\tau_{i-1}, \tau_i)$,

$$A_t = \{a_{\tau_{i-1}}\} \cup \{a_{\tau_{i-1}} + p : \exists(\theta, p, c) \in \mathcal{Y}_i, \tau_{i-1} + \theta < t < \tau_{i-1} + \theta + c\} \\ \cup \{a_{\tau_{i-1},\nu} : \nu > 1, t < \tau_{i-1} + e_{i,\nu}\}$$

$$B_t = \{b_{\tau_{i-1}}\} \cup \{b_{\tau_{i-1}} + p : \exists(\theta, p, d) \in \mathcal{Z}_i, \tau_{i-1} + \theta < t < \tau_{i-1} + \theta + d\} \\ \cup \{b_{\tau_{i-1},\nu} : \nu > 1, t < \tau_{i-1} + f_{i,\nu}\}$$

and, for any $i \in \mathbb{N}$,

$$A_{\tau_i} = \begin{cases} A_{\tau_i^-} \cup \{a_{\tau_{i-1}} + y_i\} & \text{if } \chi_i = \mathfrak{a}^- \\ A_{\tau_i^-} \setminus \{a_{\tau_{i-1}}\} & \text{if } \chi_i = \mathfrak{a}^+ \\ A_{\tau_i^-} & \text{otherwise} \end{cases}$$

$$B_{\tau_i} = \begin{cases} B_{\tau_i^-} \cup \{b_{\tau_{i-1}} + z_i\} & \text{if } \chi_i = \mathfrak{b}^+ \\ B_{\tau_i^-} \setminus \{b_{\tau_{i-1}}\} & \text{if } \chi_i = \mathfrak{b}^- \\ B_{\tau_i^-} & \text{otherwise} \end{cases}$$

where A_{t^-} is the left limit of A at t,[9] symmetrically for B.

Clearly, for the definition to be correct, it has to be

$$a_{\tau_{i-1}} + y_i > b_{\tau_{i-1}}, \qquad b_{\tau_{i-1}} + z_i < a_{\tau_{i-1}}, \qquad i \in \mathbb{N}. \tag{2}$$

3.2. Stochastic Properties of Ξ

Denote by

$$\xi_t = (a_t, b_t), \qquad t \in \mathbb{R}_+,$$

the process of the best quotes and assume that, for each $i \in \mathbb{N}$, the conditional distribution of

$$x_i \triangleq (\Delta\tau_i, \chi_i, y_i, z_i)$$

given $\bar{\xi}_{\tau_{i-1}}$ is arbitrary[10] but fulfils (2).

In the present Section, we shall postulate the stochastic properties of Ξ which, in fact, are consequences of the following two assumptions:

(i) The values of the quotes (i. e., ξ) are public information while the order books (i. e., Ξ except of ξ) are hidden.

(ii) The participants of the market interpret the history of the public data (i. e., the quotes ξ) only up to their last jump (which means that they do not take into account the time elapsed since the last jump of ξ).

To justify the assumptions, note that (i) roughly corresponds to the present reality in which the information about the quotes, including their detailed history, is easily accessible while it is difficult to obtain data of the waiting limit orders.[11] Assumption (ii), on the other hand, has no other motivation than a tractability of the further calculations. However, it may be partly justified by the fact that the changes of ξ are very frequent in practice, hence the intervals on which we require the "constant" behaviour of the participants are very short.

---

[9] Such a limit may be correctly defined by means of restrictions of A to bounded sets.

[10] We let these distributions to be "exogenous" to our model since, contrary to the limit order books, x_i may usually be observed and modelled by standard econometric techniques; hence, it is useful to assume x_i to be "brought from outside" to the model. The jumps out of the spread are, on the other hand, equal to the distance between the first and second best quotes just before the jump, i. e., they depend on the state of the order books, hence they are "endogenous" to our model.

[11] Even though they are partially published in some cases, still there is a phenomenon of hidden orders here.

Now let us translate our informal assumptions to the language of mathematics.

The fact that the participants do not interpret the time elapsed since the last jump of ξ clearly leads to their "constant" behaviour between the jumps of ξ, i. e., (conditionally) constant arrival and cancellation rates of the in-book orders and their identically (conditionally) distributed, mutually (conditionally) independent limit prices. This implies (see, e. g., [1], Lemma 6.4.VI) the conditionally Poisson (i. e., Cox) distribution of order flows and conditionally exponential lifetimes.[12]

Additionally, since the participants interpret solely the public information up to τ_{i-1} during (τ_{i-1}, τ_i), the order flow $\mathcal{Y}_i$ may depend on the history of Ξ only by means of $\bar{\xi}_{\tau_{i-1}}$ (i. e., the history of the quotes up to τ_{i-1}) and, moreover, x_i should be (conditionally) independent of $\mathcal{Y}_i$ given $\bar{\xi}_{\tau_{i-1}}$.

Clearly, analogous relations should hold for $\mathcal{Z}_i$ with an addition stemming from (i) and (ii): all the dependence between $\mathcal{Z}_i$ and the opposite order flow has to be completely explained by $\bar{\xi}_{\tau_{i-1}}$, i. e., $\mathcal{Z}_i$ and $\mathcal{Y}_i$ should be conditionally independent given $\bar{\xi}_{\tau_{i-1}}$ (we express this equivalently by including $\mathcal{Y}_i$ in the condition, see the end of the Subsection 2.2 for an explanation why).

Denote by Y_i and Z_i the ground processes of $\mathcal{Y}_i$ and $\mathcal{Z}_i$, respectively. Summarising our previous discussion, (i) and (ii) may be written as:

(α) $Y_i | \bar{\Xi}_{\tau_{i-1}}, x_i$ is a Poisson point process on $(0, \Delta\tau_i] \times \mathbb{R}_+$ with intensity $\ell \otimes \mathfrak{F}_i$, where ℓ denotes the Lebesgue measure and $\mathfrak{F}_i$ is an absolutely continuous random measure[13] on $\mathbb{R}_+$ depending only on $\bar{\xi}_{\tau_{i-1}}$.

$Z_i | \bar{\Xi}_{\tau_{i-1}}, x_i, \mathcal{Y}_i$ is Poisson on $(0, \Delta\tau_i] \times \mathbb{R}_-$ with intensity $\ell \otimes \mathfrak{G}_i$, where $\mathfrak{G}_i$ is an absolutely continuous random measure on $\mathbb{R}_-$ depending solely on $\bar{\xi}_{\tau_{i-1}}$

(β) Denoting by $(t_{i,\nu}, p_{i,\nu})_{\nu\in\mathbb{N}}$ the collection of the points of Y_i, and $(c_{i,\nu})_{\nu\in\mathbb{N}}$ their corresponding lifetimes,

$$\mathbb{P}(\cap_{j=1}^{n}[c_{i,j} > s_j] | \bar{\Xi}_{\tau_{i-1}}, x_i, Y_i) = \prod_{j=1}^{n} \exp\{-\rho_i(p_{i,j}) s_j\},$$

for any $n \in \mathbb{N}$ and $0 \le s_j < \Delta\tau_i - t_{i,j}$, $1 \le j \le n$, where ρ_i is an integrable real non-negative random function,[14] depending solely on $\bar{\xi}_{\tau_{i-1}}$ i. e.,

$$\rho_i(\bullet) = \rho(\bullet; \bar{\xi}_i)$$

for some (deterministic)

$$\rho : \mathbb{R} \otimes \mathfrak{S} \to \mathbb{R}_+$$

where $\mathfrak{S}$ is the space of truncated trajectories of ξ.

Symmetrically, if $(t_{i,\nu}, p_{i,\nu}, c_{i,\nu})_{\nu\in\mathbb{N}}$ are the points of $\mathcal{Z}_i$ then

$$\mathbb{P}(\cap_{j=1}^{n}[d_{i,j} > s_j] | \bar{\Xi}_{\tau_{i-1}}, x_i, \mathcal{Y}_i, Z_i) = \prod_{j=1}^{n} \exp\{-\sigma_i(p_{i,j}) s_j\},$$

[12] The fact that constant rates imply exponential jump times is well-known from the basic theory of Markov processes.

[13] Even if we require the absolute continuity of the limit prices here, later we show how to model the markets with discrete prices. See [3], Ch. 6, for the definition of random measure.

[14] i. e., a random element from the space L_1

for any $0 \le s_j < \Delta\tau_i - t_{i,j}, j \le n$ where $\sigma_i(\bullet) = \sigma(\bullet; \bar{\xi}_{\tau_{i-1}})$ for some measurable $\sigma : \mathbb{R}_- \times \mathfrak{S} \to \mathbb{R}_+$.

(γ) For any $n \in \mathbb{N}$ and $0 \le s_j \le \Delta\tau_i$, $2 \le j \le n$,

$$\mathbb{P}(\cap_{j=1}^n [e_{i,j} > s_j] | \bar{\Xi}_{\tau_{i-1}}, x_i, \mathcal{Y}_i, \mathcal{Z}_i) = \prod_{j=1}^n \exp\{-\rho_i(a_{\tau_{i-1},j} - a_{\tau_{i-1}})s_j\},$$

and

$$\mathbb{P}(\cap_{j=1}^n [f_{i,j} > s_j] | \bar{\Xi}_{\tau_{i-1}}, x_i, \mathcal{Y}_i, \mathcal{Z}_i, e_i) = \prod_{j=1}^n \exp\{-\sigma_i(b_{\tau_{i-1},j} - b_{\tau_{i-1}})s_j\}.$$

Having defined the stochastic properties of Ξ, let us denote, for each $i \in \mathbb{N}$, the densities of $\mathfrak{F}_i$, $\mathfrak{G}_i$ by ϕ_i and ψ_i, respectively and note that, similarly to the cancellation rates, there exist mappings $\phi : \mathbb{R}_+ \otimes \mathfrak{S} \to \mathbb{R}_+$, $\psi : \mathbb{R}_- \otimes \mathfrak{S} \to \mathbb{R}_+$ such that

$$\phi_i(\bullet) = \phi(\bullet; \bar{\xi}_{\tau_{i-1}}), \qquad \psi_i(\bullet) = \psi(\bullet; \bar{\xi}_{\tau_{i-1}}).$$

Since ϕ and ρ fully describe each $\mathcal{L}\left(\mathcal{Y}_i \left| \bar{\Xi}_{\tau_i}, x_i \right.\right)$ and since ψ and σ do the same for $\mathcal{L}\left(\mathcal{Y}_i \left| \bar{\Xi}_{\tau_i}, x_i, \mathcal{Y}_i, \mathcal{Z}_i \right.\right)$, it follows (from Disintegration Theorem, [3], 6.6.4) that the distribution of Ξ is uniquely determined by

$$\mathcal{L}\left(\Xi_0\right), \mathcal{L}\left(x_1 \left| \xi_0 \right.\right), \mathcal{L}\left(x_2 \left| \bar{\xi}_{[1]} \right.\right), \ldots$$

and by ϕ, ψ, ρ, σ (here, $\mathcal{L}(U)$ denote the distribution of U).

3.3. Discrete Version of the Model

It is surprisingly easy to get a satisfactory discrete version of our model: roughly speaking, it suffices to round the prices.

To leave an open space for describing both prices and log-prices by our model, we define the rounding function $\mathfrak{r} : \mathbb{R} \to \mathbb{R}$ generally so that

$$\mathfrak{r}(p) = \pi_\nu, \qquad p \in [\pi_\nu, \pi_{\nu+1}), \qquad \nu \in \mathbb{Z},$$

for some increasing (deterministic) sequence $(\pi_\nu)_{\nu \in \mathcal{Z}}$ and we introduce a "next tick" function $\mathfrak{n}$ defined by

$$\mathfrak{n}(\nu) = \pi_{\nu+1}, \qquad \nu \in \mathbb{Z}.$$

Denote

$$\hat{\Xi}_t = \left((\hat{A}_{t,\nu})_{\nu \in \mathbb{Z}}, (\hat{B}_{t,\nu})_{\nu \in \mathbb{Z}}\right)$$

$$\hat{A}_{t,\nu} \triangleq A_t[\pi_\nu, \pi_{\nu+1}), \qquad \hat{B}_{t,\nu} \triangleq B_t[\pi_\nu, \pi_{\nu+1})$$

and put

$$\hat{a}_t \triangleq \min\{\nu : \hat{A}_{t,\nu} > 0\}, \qquad \hat{b}_t \triangleq \max\{\nu : \hat{B}_{t,\nu} > 0\},$$

If we require $y_\bullet, z_\bullet$ to be such that

$$\hat{b}_{\tau_i} < \hat{a}_{\tau_i} \text{ for any } i \in \mathbb{N}$$

then, as it could be easily checked, $\hat{\Xi}$ follows the dynamics which we would expect from of a limit order market with discrete prices $(\pi_\nu)_{\nu \in \mathbb{Z}}$; obviously, each $\hat{A}_{\bullet,\nu}$ is interpreted as the number of waiting s.l.o.'s with the price ν; and symmetrically for $\hat{B}$.

In particular, if $\dots, \hat{\phi}_{i,-1}, \hat{\phi}_{i,0}, \hat{\phi}_{i,1}, \dots$ is a random sequence for any i and if we want the arrival rate of s.l.o.'s with the limit price π_ν between τ_{i-1} and τ_i to be $\hat{\phi}_{i,\nu}$ for each $i \in \mathbb{N}$, $\nu > \hat{a}_{\tau_{i-1}}$, then it suffices to set

$$\phi_i(p) = \sum_{\nu > \hat{a}_{\tau_{i-1}}} (\pi_{\nu+1} - \pi_\nu)^{-1} \mathbf{1}_{[\pi_\nu, \pi_{\nu+1})}(a_{\tau_{i-1}} + p) \hat{\phi}_{i,\nu}.$$

Similarly, if we require the conditional cancellation rate of the s.l.o.'s with a limit price $\nu \geq \hat{a}_{\tau_{i-1}}$ to be $\hat{\rho}_{i,\nu}$ then we may put

$$\rho_i(p) = \sum_{\nu \geq \hat{a}_{\tau_{i-1}}} \mathbf{1}_{[\pi_\nu, \pi_{\nu+1})}(a_{\tau_{i-1}} + p) \hat{\rho}_{i,\nu}.$$

Clearly, symmetric relations hold for the b.l.o.'s.

As a discrete counterpart of ξ (note that rounding has made ξ unobservable) introduce

$$\hat{X} \triangleq (\hat{a}, \hat{b}, \hat{p}, \hat{q}), \qquad \hat{p}_t \triangleq \hat{A}_{t,\hat{a}}, \quad \hat{q}_t \triangleq \hat{B}_{t,\hat{b}}, \quad t \geq 0.$$

Denote

$$0 < \hat{\tau}_0 < \hat{\tau}_1 < \dots$$

the jump times of $\hat{X}$ and note that such a jump happens if, in the underlying model, either ξ jumps, a s.l.o. with a price between a and $\mathfrak{n}(\hat{a})$ arrives or is cancelled or if a b.l.o. with a price between $\mathfrak{r}(\hat{b})$ and b arrives or is cancelled.

3.4. Basic Properties of Ξ

Finishing this Section, let us formulate several basic properties of the process Ξ, the first two guaranteeing the correctness of our definition, the third one being important in applications:

Proposition 3.1. (i) If A_0, B_0 are simple then both A_t and B_t are simple for any $t \in \mathbb{R}_+$ with probability one.

(ii) The probability that two events, including

- jumps of ξ,
- arrivals of limit orders,
- cancellations of limit orders,

happen at the same time is zero.

(iii) The process $\dot{\Xi}_t \triangleq (\Xi_t, \bar{\xi}_t)$ is Markov.

Proof. See Appendix, Section A.4. □

In the sections that follow, we shall assume, without a change of the distribution of Ξ, that the order books are simple and the events' times are exclusive not only with probability one, but for any elementary event.

4. DISTRIBUTION OF Ξ

Even if the distribution of Ξ is fully described by $\mathcal{L}(\Xi_0)$ and $(\alpha)-(\gamma)$, such a description is not very useful because it says nothing about the distribution at a (random or fixed) horizon, which is required by most of the applications. Moreover, it would be computationally demanding (for finite models) or even impossible (for infinite models) to do any Monte Carlo simulation based directly on the definition $(\alpha)-(\gamma)$. Therefore, it is very useful or even necessary to know the distribution of Ξ_θ, where $0 < \theta < \infty$ is a $\sigma(\xi_t)_{t\geq 0}$-optional time (i. e., a random variable for which it suffices to know $\bar{\xi}_t$ in order to decide whether $\theta \leq t$ for any $t \in \mathbb{R}$. See [3], Chapter 7, for the definition and basic properties of optional times); we describe this distribution forthwith.

We divide our task into two steps: the description of $\mathcal{L}\left(\Xi_\theta \left|\bar{\xi}_\theta\right.\right)$ (Subsection 4.1) and a (recursive) expression of the distribution of $\mathcal{L}(\xi)$ (hence of $\mathcal{L}\left(\bar{\xi}_\theta\right)$, Subsection 4.2).

4.1. Distribution of $\Xi|\xi$

Let us proceed to the conditional distribution of Ξ_θ. Due to the symmetry of our model, only the distribution of A_θ will be treated in full detail here (the distribution of B_θ may be obtained by mirroring); in particular, we shall examine the conditional distribution of A_θ given the initial state of the market Ξ_0 and the history of the best quotes $\bar{\xi}_\theta$.

In the beginning, we will make a preliminary analysis of the composition of A_θ; let us start by noting that an order with an arrival time t and a limit price p is unexecuted at a time θ if and only if its limit price is not smaller than the best ask since its appearance, i. e.,

$$p \geq \tilde{a}_{I(t)}, \qquad I(t) \triangleq \max\{i : \tau_i \in [0,t]\}, \tag{3}$$

where

$$\tilde{a}_i = \tilde{a}_i^\theta \triangleq \max_{i\leq\nu\leq\tilde{I}} a_{\tau_\nu}, \qquad 0 \leq i \leq \tilde{I}, \qquad \tilde{I} \triangleq I(\theta), \text{[15]}$$

see Figure 1 for a graphical illustration. Hence, by a pure logic, the p.p. A_θ, being a set of all the uncancelled and unexecuted limit orders having arrived until θ, may be split into the following four disjoint p.p.'s:

$\{a_\theta\}$ — the present ask

[15] *Proof.* When the order is unexecuted, (3) has to hold, otherwise there would exist an unexecuted order with a limit price smaller than the best ask. Conversely, if an order fulfilling (3) were executed then, due to the fact that A is simple, its execution would cause a jump of a above its limit price at the time of the execution, which would contradict (3).

D — the set of the uncancelled initial limit orders (i. e., those which were present in A_0) fulfilling (3) with ">"

L — the set of all the uncancelled limit orders with a *positive* arrival time fulfilling (3) with ">"

E — the set of all the uncancelled orders (initial or with a positive arrival time) fulfilling (3) with "=" but differing from a_θ.

Proceeding further in the analysis, note that

$$L = L_1 + \cdots + L_{\tilde{I}+1}$$

where, for each $1 \leq i \leq \tilde{I}+1$, L_i is the p.p. of (the limit prices of) the uncancelled orders with arrival times belonging to $[\tau_{i-1}, \tau_i \wedge \theta)$ and with limit prices being strictly larger than $\tilde{a}_{i-1}$, i. e., those l.o.'s whose arrival time and limit price falls into the rectangle

$$\Lambda_i \triangleq [\tilde{\tau}_{i-1}, \tilde{\tau}_i) \times (\tilde{a}_{i-1}, \infty)$$

where

$$\tilde{\tau}_i = \tilde{\tau}_i^\theta \triangleq \begin{cases} \tau_i & \text{if } i \leq \tilde{I}, \\ \theta & \text{if } i = \tilde{I}+1, \end{cases}$$

(see Figure 1 for an illustration).

The following Proposition describes the conditional distributions of the individual components of an order book and proves their mutual (conditional) independence.

Proposition 4.1. (conditional distribution of order books) Let θ be a finite positive $\sigma(\xi_t)$-optional time. For any i, denote

$$f_i(p) = f_i^\theta(p) \triangleq \begin{cases} \phi_i(p)\Delta\tilde{\tau}_i & \text{if } \rho_i(p) = 0, \\ \frac{\phi_i(p)}{\rho_i(p)}[1 - e^{-\Delta\tilde{\tau}_i\rho_i(p)}] & \text{if } \rho_i(p) > 0, \end{cases}$$

$$\epsilon(i,x) = \epsilon^\theta(i,x) \triangleq \exp\left\{-\sum\nolimits_{j=i}^{I(\theta)} \rho_{j+1}(x - a_{\tau_j})\Delta\tilde{\tau}_{j+1}\right\}$$

and

$$\dot{A}_i = \dot{A}_i^\theta \triangleq A_{\tilde{\tau}_i}|_{(\tilde{a}_i,\infty)}.$$

Further, denote by

$$\breve{A} = \breve{A}^\theta \triangleq [\breve{a}_1, \dots, \breve{a}_J]$$

the collection of all the distinct elements of $\{\tilde{a}_0, \dots, \tilde{a}_{\tilde{I}}\} \setminus \{a_\theta\}$ (i. e., the set of the former asks which could potentially survive until θ) and

$$\gamma_\nu = \gamma_\nu^\theta = \min\{k : \tilde{a}_k < \breve{a}_\nu\}, \qquad 1 \leq \nu \leq J,$$

(i. e., the index of the time when $\breve{a}_\nu$ ceased to be an ask). Then it holds that

($\circ$) $A_\theta = \{a_\theta\} + L + D + E$,

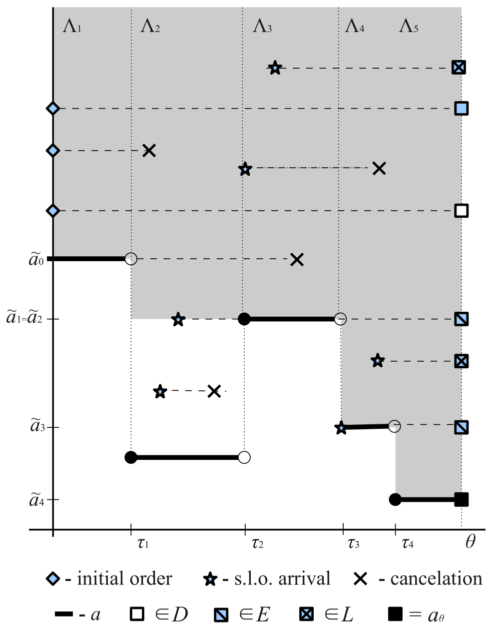

Fig. 1. Composition of the order book.

where, denoting $C = C^\theta = (\Xi_0, \theta, \bar{\xi}_\theta)$,

(−) a_θ is conditionally constant given C,

(i) $L|C$ is a Poisson point process with density

$$F(p) \triangleq \sum_{i=1}^{\tilde{I}+1} \mathbf{1}_{(\tilde{a}_{i-1},\infty)}(p) f_i(p - a_{\tau_{i-1}})\epsilon(i,p),$$

(ii) $D|C$ is an h-thinning of $\dot{A}_0 = [\dot{a}_1, \dot{a}_1, \dots]$ with parameters

$$\epsilon(0,\dot{a}_1), \epsilon(0,\dot{a}_2), \dots,$$

(iii) $E|C$ is an h-thinning of $\breve{A} = [\breve{a}_1, \dots, \breve{a}_J]$ with parameters

$$\epsilon(\gamma_1, \breve{a}_1), \dots, \epsilon(\gamma_J, \breve{a}_J),$$

(iv) L, D, E are mutually conditionally independent given C.

If, in addition, A_0 is a Poisson point process with density ϕ_0 then all (∘), (−), (i), (iii), (iv) hold true with $C' \triangleq (\theta, \bar{\xi}_\theta)$ instead of C and (ii') $D|C'$ is a Poisson process with density

$$F_0(p) = \mathbf{1}_{(\tilde{a}_0,\infty)}(p)\phi_0(p)\epsilon(0,p).$$

Moreover,

($*$) formulas symmetric to ($\circ$), ($-$), (i)$-$(iv), (ii') hold for B

and

($\triangle$) A_θ is conditionally independent of B_θ both given C and given C'.

P r o o f. The proof may be found in Appendix A.4. $\square$

4.2. Distribution of ξ

Let us proceed to the distribution of ξ; since the distributions of $x_1, x_2, \dots$ have been in fact postulated, it only remains to add a formula for the distribution of the quote jumps out of the spread.

Proposition 4.2. (Distribution of ξ) For any $k \in \mathbb{N}$, denote by

$$C_k \triangleq (\Xi_0, \bar{\xi}_{\tau_{k-1}}, \Delta\tau_k, \chi_k).$$

(i) The distribution a_{τ_k} given C_k is determined as follows: For any $p < 0$,

$$\mathbb{P}[\Delta a_{\tau_k} \le p|C_k] = \begin{cases} \mathbb{P}[y_k < p|C_k] & \text{if } \chi_k = \mathfrak{a}^- \\ 0 & \text{if } \chi_k \neq \mathfrak{a}^- \end{cases} \tag{4}$$

while, for any $p \ge 0$,

$$\mathbb{P}[\Delta a_{\tau_k} > p|C_k] = \begin{cases} 0 & \text{if } \chi_k \neq \mathfrak{a}^+ \\ H_k(p) & \text{if } \chi_k = \mathfrak{a}^+ \end{cases} \tag{5}$$

where

$$H_k(p) \triangleq \exp\left\{-\int_{-\infty}^{a_{\tau_{k-1}}+p} \grave{F}(z)\,\mathrm{d}z\right\} \cdot \prod_{\alpha \in \dot{A}_0^{\tau_{k-1}}, \alpha \le a_{\tau_{k-1}}+p} [1-\epsilon^{\tau_k}(0,\alpha)]$$
$$\cdot \prod_{1 \le \nu \le |\breve{A}^{\tau_{k-1}}|, \breve{a}_\nu^{\tau_{k-1}} \le a_{\tau_{k-1}}+p} [1-\epsilon^{\tau_k}(\gamma_\nu^{\tau_{k-1}}, \breve{a}_\nu^{\tau_{k-1}})],$$

(ϵ^{τ_k} stands for ϵ^{τ_k}, $\breve{A}^{\tau_{k-1}}$ for $\breve{A}^{\tau_{k-1}}$, etc.), and where

$$\grave{F}(p) \triangleq \sum_{\nu=1}^{k} \mathbf{1}_{(\tilde{a}_{\nu-1}^{\tau_{k-1}},\infty)}(p) f_\nu^{\tau_k}(p - a_{\tau_\nu})\epsilon^{\tau_k}(\nu,p).$$

(i') If A_0 is a P.p.p. with density ϕ_0 and $C'_k \overset{\triangle}{=} (\bar{\xi}_{\tau_{k-1}}, \Delta\tau_k, \chi_k)$ then (4) holds with C'_k instead of C_k while, for any $p \geq 0$,

$$\mathbb{P}[\Delta a_{\tau_k} > p | C'_k] = \begin{cases} 0 & \text{if } \chi_k \neq \mathfrak{a}^+ \\ H'_k(p) & \text{if } \chi_k = \mathfrak{a}^+ \end{cases}$$

where

$$H'_k(p) \overset{\triangle}{=} \exp\left\{ - \int_{-\infty}^{a_{\tau_{k-1}}+p} [\grave{F}(z) + \phi_0(z) \epsilon^{\tau_k}(0,z)] \,\mathrm{d}z \right\} \cdot \prod_{\alpha \in \dot{A}_0^{\tau_k - 1}, \alpha \leq a_{\tau_{k-1}}+p} [1 - \epsilon^{\tau_k}(0,\alpha)],$$

($*$) formulas symmetric to (i) and (i') hold for b,

($\triangle$) a_{τ_k} and b_{τ_k} are conditionally independent both given C_k and given C'_k.

Proof. See Appendix, A.4. □

Remark 4.3. Note that Propositions 4.1 and 4.2 fully describe $\mathcal{L}\left(\dot{\Xi}_\tau \middle| \dot{\Xi}_0\right)$ for any $\theta \in \mathbb{R}_+$; therefore and by Proposition 3.1 (iii), we have fully characterised the distribution of $\dot{\Xi}$, hence also Ξ.

5. CONCLUSION

In our paper, a general model of the continuous double auction was formulated and its distribution was specified.

Our model provides a unified framework for the description of three existing models of the continuous double auction ([4, 5] and [11]) enabling, among other things, to test those (and similar) models statistically or to construct efficient simulation techniques (see e. g. Section 6 of [10]).

To conclude, let us stress that we could proceed much further in the direction taken in this paper. For instance, the assumption of unit order sizes could be relaxed, at least for the continuous models (the only changes would be that the Poisson variables would become compound Poisson and that the formulas for the distributions of a and b would become complicated). However, since further additions would considerably increase the size of the present introductory paper, we leave these and many other possibilities for our or someone else's future research.

A. APPENDIX

The Appendix is organized as follows: first, the relation of our setting to the covered models is described, second, auxiliary results are formulated, third, the model is reformulated as a mapping of mutually independent random elements (this reformulation is further used in the proofs). Finally, the proofs of the Propositions from the main text are presented.

A.1. Link to Existing Models

As it was written in the Introduction, three existing models, namely those of [4, 5] and [11] are special cases of our model. In the present Subection, we clarify the relations of those models to our one, in particular, we re-define them using our notation. A more detailed treatment of this topic may be found in [10].

Maslov's Model

The model by [5], being the earliest of the mentioned ones, assumes i.i.d. (exponential) inter-event times where the events include arrivals of either buy or sell, limit or market orders (there are no cancellations in the model). The probability that the newly arriving order is of any particular type is 1/4 and the limit price of a newly arriving limit order is unit uniform on the interval starting at the opposite best quote. Implicitly, it is assumed that the inter-event times, the selection of a type of a newly coming order and the limit prices are mutually independent.

If we agree to consider an arrival of a market order at a time when the opposite order book is empty as a jump of $\bar{\xi}$ too, then the ith jump of ξ, i. e. the arrival of either a market or an in-spread limit order, satisfies

$$\mathcal{L}\left(\Delta\tau_i|\xi_{\tau_{i-1}}\right) = \mathrm{Exp}\left(s_i/4\right), \qquad s_i = 2[(a_{\tau_{i-1}} - b_{\tau_{i-1}}) \wedge 1] + 2,$$

while the remaining characteristics are:

$$\mathbb{P}[\chi_i = \mathfrak{a}^+|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \mathbb{P}[\chi_i = \mathfrak{b}^-|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{1}{s_i}$$

$$\mathbb{P}[\chi_i = \mathfrak{a}^-|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \mathbb{P}[\chi_i = \mathfrak{b}^+|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{(a_{\tau_{i-1}} - b_{\tau_{i-1}}) \wedge 1}{s_i}$$

$$\mathcal{L}\left(-y_i|\chi_i = \mathfrak{a}^-, \tau_i, \bar{\xi}_{\tau_{i-1}}\right) = \mathcal{L}\left(z_i|\chi_i = \mathfrak{b}^+, \tau_i, \bar{\xi}_{\tau_{i-1}}\right) = \mathrm{U}\left(0, (a_{\tau_{i-1}} - b_{\tau_{i-1}}) \wedge 1\right)$$

$$\phi_i(p) = \psi_i(-p) = \frac{1}{4}\mathbf{1}_{(0,[1-(a_{\tau_{i-1}}-b_{\tau_{i-1}})]\vee 0)}(p), \qquad \rho_i \equiv \sigma_i \equiv 0$$

where $\mathbf{1}$ denotes the characteristic function.

Luckock's Model

Another finite model, introduced by [4], assumes unit inter-event times, no cancellations, equal probability that an incoming agent is a buyer or a seller, and i.i.d. buying and selling absolute *reservation* (i. e. intended) *prices*. The newly coming agent takes his actions as follows: if he is a buyer and his reservation price is less than the ask then he puts a limit order with the limit price equal to the reservation price, otherwise he immediately buys (by putting a market order), the situation for sellers is symmetric. Hence, the only parameters of the Luckock's model are the (continuous) distribution functions of the selling and buying reservation prices, denoted by K, L, respectively.

In the language of our model, the model is defined as

$$\mathcal{L}\left(\Delta\tau_i|\bar{\xi}_{\tau_{i-1}}\right) = \mathrm{Exp}\left(S_i/2\right), \qquad S_i = K(a_{\tau_{i-1}}) + [1 - L(b_{\tau_{i-1}})],$$

$$P[\chi_i = \mathfrak{b}^+|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{L(a_{\tau_{i-1}}) - L(b_{\tau_{i-1}})}{S_i}$$

$$\mathbb{P}[\chi_i = \mathfrak{a}^-|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{K(a_{\tau_{i-1}}) - K(b_{\tau_{i-1}})}{S_i}$$

$$\mathbb{P}[\chi_i = \mathfrak{a}^+|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{1 - L(a_{\tau_{i-1}})}{S_i} \qquad \mathbb{P}[\chi_i = \mathfrak{b}^-|\tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{K(b_{\tau_{i-1}})}{S_i}$$

$$\mathbb{P}[y_i < p|\chi_i = \mathfrak{a}^-, \tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{K(a_{\tau_{i-1}} + p) - K(b_{\tau_{i-1}})}{K(a_{\tau_{i-1}}) - K(b_{\tau_{i-1}})}$$

$$\mathbb{P}[z_i < p|\chi_i = \mathfrak{b}^+, \tau_i, \bar{\xi}_{\tau_{i-1}}] = \frac{L(b_{\tau_{i-1}} + p) - L(b_{\tau_{i-1}})}{L(a_{\tau_{i-1}}) - L(b_{\tau_{i-1}})}$$

$$\phi_i(p) = \frac{1}{2}L'(a_{\tau_{i-1}} + p) \qquad \psi_i(p) = \frac{1}{2}K'(b_{\tau_{i-1}} + p) \qquad \rho_i \equiv \sigma_i \equiv 0$$

where $'$ means a derivative.

Smith's and Farmer's Model

The probably most popular model of the c.d.a. is the infinite one by [11]. It assumes unit ticks, a constant rate η of arrivals of both the types of market orders, a constant per tick arrival rate ς of both the types of limit orders and a constant rate ι of cancellations.

In the language of our (general) definition, this means that $\mathfrak{r}(p) = \lfloor p \rfloor$ (i. e. the integer part of p) and, for each $i \in \mathbb{N}$,

$$\Delta\tau_i|\bar{\xi}_{[i-1]} \sim \operatorname{Exp}(\omega_i), \qquad \omega_i = 2(\eta + \iota) + [s_i^a + s_i^b]\varsigma,$$

$$\chi_i|\bar{\xi}_{[i-1]} = \left\{ \frac{\eta + \iota}{\omega_i}, \frac{s_i^a\varsigma}{\omega_i}, \frac{s_i^b\varsigma}{\omega_i}, \frac{\eta + \iota}{\omega_i} \right\}$$

where

$$s_i^a = a_{\tau_{i-1}} - (\hat{b}_{\tau_{i-1}} + 1), \qquad s_i^b = \hat{a}_{\tau_{i-1}} - b_{\tau_{i-1}},$$

and, further,

$$y_i|\bar{\xi}_{[i-1]} \sim \mathrm{U}(-s_i^a, 0), \qquad \phi_i \equiv \varsigma, \qquad \rho_i \equiv \iota,$$

$$z_i|\bar{\xi}_{[i-1]} \sim \mathrm{U}\left(0, s_i^b\right), \qquad \psi_i \equiv \varsigma, \qquad \sigma_i \equiv \iota.$$

Since, for any $i \in \mathbb{N}$, $\Delta\tau_i$ is conditionally exponential, it is easy to determine the distribution of $\Delta\hat{\tau}_i$ (see Subsection 3.3 for a definition) since the first jump of $\hat{X}$ after $\hat{\tau}_{i-1}$ happens if and only if, in the underlying model, one of the following events happen:

- ξ jumps (which happens with rate ω_j where j is the index of the last $\tau_\bullet$) or
- an order arrives between a and $\hat{a}$ (with rate $(a - \hat{a})\varsigma$) or
- an order (except the ask) with price between $\hat{a}$ and $\hat{a} + 1$ is cancelled (with rate $(\hat{p} - 1)\iota$) or

- an order arrives between b and $\hat{b}+1$ (with rate $(\hat{b}+1-b)\varsigma$) or
- an order (except the bid) with price between $\hat{b}$ and $\hat{b}+1$ is cancelled (with rate $(\hat{q}-1)\iota$),

which, summarized, gives

$$\Delta\hat{\tau}_i|\hat{\bar{\Xi}}_{\hat{\tau}_{i-1}} \sim \mathrm{Exp}\left(\hat{\omega}_i\right), \tag{6}$$

where

$$\hat{\omega} \stackrel{\triangle}{=} 2\eta + (\hat{p}_{\hat{\tau}_{i-1}} + \hat{q}_{\hat{\tau}_{i-1}})\iota + (\hat{a}_{\hat{\tau}_{i-1}} - \hat{b}_{\hat{\tau}_{i-1}})\varsigma.$$

A.2. Lemmas and Auxiliary Results

Before starting to deal with auxiliary results and proofs, let us agree to assume that all the underlying random elements (12) as well as eventual further randomisations are defined on a common probability space $(\Omega, \mathbb{P}, \mathcal{F})$, which is rich enough.

Further, let us introduce an additional notation: By writing $X \oplus_i c$ or $X \ominus_i c$, we shall mean a shift of a point process (or a measure) X in the ith coordinate by c to the left or to the right, respectively. If, for instance, X is a p.p. on $\mathbb{R}^3$ with points $(x_{i,1}, x_{i,2}, x_{i,3})_{i\in\mathbb{N}}$ then the points of $X \ominus_2 c$ are $(x_{i,1}, x_{i,2} - c, x_{i,3})_{i\in\mathbb{N}}$. For $n = 1$ we shall omit the index at $\oplus$ or $\ominus$.

The notion $\stackrel{d}{=}$ means equality of distributions.

Lemma A.1. (distribution of functions of independent elements) Let U and V be mutually independent random elements taking values in measurable spaces $(L, \mathcal{L})$ and $(M, \mathcal{M})$, respectively. Let $(Q, \mathcal{Q})$ be a measurable space and let $f : L \times M \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ and $F : L \times M \to (Q, \mathcal{Q})$ be measurable mappings. Then

(i) $\mathbb{E}f(U, V) = \mathbb{E}g(V)$, where $g(v) = \mathbb{E}f(U, v), v \in M$.

(ii) If, for each $v \in M$ and $z \in Q$,

$$\mathbb{P}(f(U, v) \in \bullet | F(U, v) = z) = p_z(\bullet) \tag{7}$$

where $p_\bullet$ is a conditional probability distribution (viewed as a probability measure dependent on the condition) then

$$\mathbb{P}(f(U, V) \in \bullet | F(U, V) = z) = p_z(\bullet). \tag{8}$$

Violating the strict formality but sparing much notation, we may compute expectations or conditional distributions of the type in (i), (ii) respectively, as if V was deterministic and use Lemma A.1 to obtain the validity of our computations given the true distribution of (U, V).

P r o o f. For the proof of (i), see [2], 4.5.2.

(ii) For any $A \in \mathfrak{B}(\mathbb{R})$ and $B \in \mathcal{Q}$,

$$\begin{aligned}
&\mathbb{P}(f(U,V) \in A, F(U,V) \in B)\\
&= \int\int \mathbf{1}_A(f(u,v))\mathbf{1}_B(F(u,v))\,\mathrm{d}\mathbb{P}_U(u)\mathrm{d}\mathbb{P}_V(v)\\
&= \int P(f(U,v) \in A, F(U,v) \in B)\,\mathrm{d}\mathbb{P}_V(v)\\
&= \int \left(\int_B P(f(U,v) \in A | F(U,v) = z)\,\mathrm{d}\mathbb{P}_{F(U,v)}(z)\right)\mathrm{d}\mathbb{P}_V(v)\\
&= \int \left(\int_B p_z(A)\,\mathrm{d}\mathbb{P}_{F(U,v)}(z)\right)\mathrm{d}\mathbb{P}_V(v)\\
&= \int \left(\int \mathbf{1}_B(z)p_z(A)\,\mathrm{d}\mathbb{P}_{F(U,v)}(z)\right)\mathrm{d}\mathbb{P}_V(v)\\
&= \int \left(\int \mathbf{1}_B(F(u,v))p_{F(u,v)}(A)\,\mathrm{d}\mathbb{P}_U(u)\right)\mathrm{d}\mathbb{P}_V(v)\\
&= \int \mathbf{1}_B(z)p_z(A)\,\mathrm{d}\mathbb{P}_{F(U,V)}(z) = \int_B p_z(A)\,\mathrm{d}\mathbb{P}_{F(U,V)}(z)
\end{aligned}$$

which proves (8) by the definition of the conditional probability, viewed as a function of the condition. □

The next assertion will be found as helpful as the previous one:

Lemma A.2. (Local property) Let $u, \dot{u}$ be real random variables and let $S \subset \mathcal{F}$ such that $u = \dot{u}$ on S.

(i) If $\mathcal{G} \subseteq \mathcal{F}$, $\mathcal{H} \subseteq \mathcal{F}$ are such that $S \in \mathcal{G}$, $S \in \mathcal{H}$ and $\mathcal{G}|_S = \mathcal{H}|_S$ then $\mathbb{E}(u|\mathcal{G}) = \mathbb{E}(\dot{u}|\mathcal{H})$ w.p.1 on S.

(ii) If, in addition, U and $\dot{U}$ are random elements taking values in a measurable space $(K, \mathcal{K})$ such that $S \in \sigma(U)$, $S \in \sigma(\dot{U})$ and $U = \dot{U}$ on S then $\mathbb{E}(u|U) = \mathbb{E}(\dot{u}|\dot{U})$ w.p.1 on S.

Proof. (i) [3], Lemma 6.2.
(ii) Since $\sigma(U) \cap S = U^{-1}(\mathcal{K}) \cap S = \dot{U}^{-1}(\mathcal{K}) \cap S = \sigma(\dot{U}) \cap S$, i.e., $\sigma(U)$ and $\sigma(\dot{U})$ coincide on S, the assertion follows from (i). □

Our usual usage of Lemma A.2 follows: in attempt to prove a (complicated) relation concerning conditional expectations (or distributions), we simplify our situation by partitioning Ω into sets $S_1, S_2, \dots \in \mathcal{F}$ on which it is possible to prove the relation by means of Lemma A.2; the general validity of the relation will follow from the additive property of conditional expectations (using the fact that there are only countably many sets $S_\bullet$).

Lemma A.3. (Representation by uniform random elements) For any non-decreasing right continuous function G defined on $\mathbb{R}$ with $G(-\infty^+) = 0$, denote

$$G^{[-1]}(p) = \inf\{x : G(x) \geq p\}, \qquad p \in (0, G(\infty^-)), \tag{9}$$

its generalised inversion.

The following assertions hold:

(*) If G is continuous then $G^{[-1]}(\bullet)$ is a bijection between $(0, G(\infty^-))$ and $T_G \stackrel{\triangle}{=} G^{[-1]}((0, G(\infty^-)))$.

(i) If z is a k-dimensional real random vector and u is a k-dimensional uniform random vector on the unit cube $(0,1)^k$ then

$$z \stackrel{d}{=} \gamma(u),$$

$$\gamma : (0,1)^k \to \mathbb{R}^k, \qquad \gamma_1(p) = G_1^{[-1]}(p_1),$$
$$\gamma_i(u) = G_i^{[-1]}(u_i|\gamma_1(u), \ldots, \gamma_{i-1}(u)), \qquad 2 \leq i \leq n,$$

where G_1 is the cumulative distribution function (c.d.f.) of the first component of z and where $G_i(\bullet|\epsilon_1, \ldots, \epsilon_{i-1})$ is the conditional c.d.f. of the ith component of z given its first $i-1$ components for each $i \leq k$.

(ii) If Z is a marked P.p.p on $\mathbb{R}_+$ with absolutely continuous intensity μ and independent marks (in the sense of [1], ch. 6.4.) taking values in $\mathbb{R}^{k-1}$ for some $k \in \mathbb{N}$ and if U is a unit P.p.p. on $(0, \mu(\mathbb{R}_+)) \times (0,1)^{k-1}$ then

$$Z \stackrel{d}{=} \Gamma(U),$$

$$\Gamma_1(u) = G^{[-1]}(u_1), \qquad G(x) \stackrel{\triangle}{=} \mu(0, x),$$
$$\Gamma_i(u) = G_{i-1}^{[-1]}(u_i|\Gamma_1(u), \ldots, \Gamma_{i-1}(u)), \qquad 2 \leq i \leq k,$$

where $G_\nu(\bullet]\zeta, m_1, \ldots, m_{\nu-1})$ is the conditional c.d.f. of the νth component of the mark given the location of the corresponding point ζ and the first $\nu-1$ components of the marks, $\nu \leq k-1$.

(iii) If Z is the same as in (ii) and if V is P.p.p. on $\mathbb{R}_+ \times [0,1]^{k-1}$ with intensity $\mu \otimes \ell^{k-1}$ then $Z \stackrel{d}{=} \Upsilon(V)$ where

$$\Upsilon_1(u) = u_1,$$
$$\Upsilon_i(u) = G_{i-1}^{[-1]}(u_i|u_1, \Upsilon_2(u) \ldots, \Upsilon_{i-1}(u)), \qquad 2 \leq i \leq n.$$

This Lemma may be useful in two ways: in Monte Carlo simulations (the computers usually generate only uniform random elements) and in theoretical constructions, used for representations of various random objects by means of "distribution-free" underlying elements.

Proof. ($\star$) It is easy to check that if $G^{[-1]}$ was not a bijection then G would jump.

(i) For the case $k = 1$, see, e. g., [7], p. 238. For the general case, see [8], Lemma 1.12.

(ii) Denote by Z_g the ground process of Z, by Z'_g the ground process of $\Gamma(U)$, and by U_g the ground process of U viewed as a marked real p.p. (note that the marks of such a process are i.i.d. uniform on $(0,1)^{k-1}$ by [1], Lemma 6.4.VI). Let us start by proving that

$$Z_g \stackrel{d}{=} Z'_g. \tag{10}$$

Since, for any $S = (-\infty, a]$, $a \in \mathbb{R}$,

$$\mu S = G(a) = \ell(0, G(a)] = \ell G(S)$$

and since half lines generate $\mathfrak{B}(\mathbb{R})$, we are getting

$$\mu S = \ell G(S), \qquad S \in \mathfrak{B}(\mathbb{R}), \tag{11}$$

by the monotone class argument [3, Theorem 1.1.].

Due to $(\star)$ and since we may assume that no points of Z'_g lie in $\mathbb{R}_+ \setminus T_G$[16] it holds that s is a point of Z'_g if and only if $G(s)$ is a point of U_g which further yields

$$Z'_g S = U_g(G(S)), \qquad S \in \mathfrak{B}(\mathbb{R}),$$

further implying

$$\mathcal{L}\left(Z'_g S\right) = \mathcal{L}\left(U_g(G(S))\right) = \text{Poisson}\left(\ell(G(S))\right) \stackrel{(11)}{=} \text{Poisson}\left(\mu S\right), \qquad S \in \mathfrak{B}(\mathbb{R}),$$

which is sufficient for (10) by the uniqueness criterion for simple point processes [3, Theorem 12.8].

Now, note that if

$$(\spadesuit) \qquad U_g \text{ was deterministic,}$$

then the conditional distribution of the marks given the locations of the points would become unconditional; hence we could use assertion (i) of the present Lemma to prove that the distribution of marks (given $(\spadesuit)$) of Z' is the same as the one determined by $G_1, \dots, G_{k-1}$ (with fixed ζ). Moreover, since the marks of U are independent of U_g (by [1], Lemma 6.4.VI), we may use assertion (ii) of Lemma A.1 to get that the same is true given the true distribution of U_g which, in combination with (10), proves the assertion of the Lemma by the Disintegration Theorem ([3], 6.6.4).

(iii) The ground processes of Z and $\Upsilon(U)$ agree in distribution trivially, the proof of equality of the distributions of the marks is the same as in assertion (ii). □

[16] Proof: Due to the Lebesgue decomposition of G, the set $T' \stackrel{\triangle}{=} \mathbb{R}_+ \setminus T_G$ consists of countably many intervals. Hence, if $\mu(T') > 0$ then there would exist an interval $I \subseteq T'$ on which G would be strictly increasing which would contradict the definition of T_G. Therefore, $\mu(T') = 0$ so a possible change of Z'_g on T' would not change its distribution.

A.3. Ξ as a Function of Independent Elements

Even if we have fully characterised the distribution of Ξ in Subsections 3.1 and 3.2, it will be much simpler to handle the model if we express Ξ (without the change of its distribution) as a function of mutually independent random elements; we find this useful especially during the derivation of the distribution of Ξ (Appendix A.4).

For any $i \in \mathbb{N}$, let $\mathcal{V}_i$ and $\mathcal{W}_i$ be P.p.p.'s on $\mathbb{R}_+ \times \mathbb{R}_+ \times (0,1)$ with the unit intensity, let $v_{i,2}, w_{i,2}, v_{i,3}, w_{i,3}, \dots$ be unit exponential random variables and let $u_i \in (0,1) \times (0,1) \times (0,1)$ be a three dimensional-uniform random vector. Further, let all

$$\begin{array}{c} \mathcal{V}_1, \mathcal{W}_1, u_1, v_{1,2}, w_{1,2}, v_{1.3}, v_{1.3}, \dots \\ \mathcal{V}_2, \mathcal{W}_2, u_2, v_{2,2}, w_{2,2}, v_{2.3}, v_{2.3}, \dots \\ \dots \end{array} \tag{12}$$

be mutually independent and independent of Ξ_0.

Starting to (re)construct Ξ, note that there exists a mapping F such that, for any $i \in \mathbb{N}$, the conditional distribution of x_i given $\bar{\xi}_{\tau_{i-1}}$ will not change if, for each $i \in \mathbb{N}$,

$$x_i = F(u_i; \bar{\xi}_{\tau_{i-1}}) \tag{13}$$

(to see it, assume for a while that $\bar{\xi}_{\tau_{i-1}}$ is deterministic and construct F by means of conditional c.d.f.'s as in Lemma A.3 (i) to get

$$F(u_i; \bar{\xi}_{\tau_{i-1}}) \overset{d}{=} x_i \tag{14}$$

and note that, by Lemma A.1 (ii), formula (14) holds also for the true distribution of $\bar{\Xi}_{\tau_{i-1}}$).

Further, denote by $\Phi(\bullet;\bullet)$ and $\Psi(\bullet;\bullet)$ the c.d.f.'s corresponding to ϕ, ψ, respectively, and, for each $i \in \mathbb{N}$, put

$$\mathcal{Y}_i = G(\mathcal{V}_i|_{(0,\Delta\tau_i)\times(0,\dot{v}_i)\times(0,1)}; \bar{\xi}_{\tau_{i-1}}), \tag{15}$$

$$\mathcal{Z}_i = H(\mathcal{W}_i|_{(0,\Delta\tau_i]\times(0,\dot{w}_i)\times(0,1)}; \bar{\xi}_{\tau_{i-1}}), \tag{16}$$

where $\dot{v}_i \triangleq \Phi(\infty^-; \bar{\xi}_{\tau_{i-1}})$, $\dot{w}_i \triangleq \Psi(0; \bar{\xi}_{\tau_{i-1}})$,

$$G(t,\pi,u;\zeta) \to \left(t, \Phi^{[-1]}(\pi;\zeta), \frac{u}{\rho(\Phi^{[-1]}(\pi;\zeta);\zeta)}\right)$$

and where H is defined symmetrically (see Lemma A.3 for a definition of $\bullet^{[-1]}$).

By Lemma A.1 (i) (with $V = \bar{\xi}_{\tau_{i-1}}$) and Lemma A.3 (ii) (taking the second coordinate of $\mathcal{Y}_i$ as the ground process and the first one together with the third one as a mark; note that the first coordinates are i.i.d. $\sim \mathrm{U}(0,\Delta\tau_i)$), we get that $\mathcal{Y}_i$ and $\mathcal{Z}_i$ defined by (15) and (16) fulfil (α) and (β), respectively, including the required conditional independences (the latter stemming from the independence of underlying variables).

To finish the construction, put, for each $i \in \mathbb{N}$,

$$e_{i,\nu} = \begin{cases} \frac{v_{i,\nu}}{\rho(a_{\tau_{i-1},\nu} - a_{\tau_{i-1}}; \bar{\xi}_{\tau_{i-1}})} & \text{if } a_{\tau_{i-1},\nu} < \infty \text{ and } \rho(a_{\tau_{i-1},\nu} - a_{\tau_{i-1}}; \bar{\xi}_{\tau_{i-1}}) \neq 0 \\ \infty & \text{otherwise} \end{cases} \tag{17}$$

and define $f_\bullet$ symmetrically by means of $w_\bullet$. By assertion (ii) of Lemma A.1 with temporarily deterministic $\bar{\Xi}_{\tau_{i-1}}$, x_i, Y_i and Z_i we immediately get that our construction fulfils (γ) as well.

A.4. Proofs

The present Section contains proofs of the Propositions stated in the main text.

Proof of Proposition 3.1. (i) Let $i \in \mathbb{N}$ and assume that $A_{\tau_{i-1}}$ and $B_{\tau_{i-1}}$ are simple. If

$(\heartsuit_i)$ the elements Ξ_0 and $(u_\nu, \mathcal{V}_\nu, \mathcal{W}_\nu, v_{\nu,\bullet}, w_{\nu,\bullet})_{\nu<i}$ are deterministic

then, thanks to the continuity of $\mathfrak{F}_i$ and $\mathfrak{G}_i$,

$$\mathbb{P}[\text{Two orders share the same price during } (\tau_{i-1}, \tau_i]] = 0.$$

Since, by assertion (i) of Lemma A.1, the same is true for the actual distribution of the variables from $(\heartsuit_i)$, point (i) of the present Proposition is proven for $t \in [\tau_{i-1}, \tau_i)$. By induction, we get (i) for the whole $\mathbb{R}_+$.
The proof of (ii) is similar: if $(\heartsuit_i)$ held and, in addition, u_i was deterministic then, by the continuity of the Lebesgue measure,

$$\mathbb{P}[\text{Two events share the same time during } (\tau_{i-1}, \tau_i]] = 0;$$

hence the general validity of this follows analogously to (i).

Before proceeding to (iii), note that, for each $i \in \mathbb{N}$ and $\nu \geq 2$, $v_{i,\nu}$ may be regarded as the first jump of a unit P.p.p. on $\mathbb{R}_+$; we will denote the process by $v^\star_{i,\nu}$ (similarly with $w_{i,\nu}$ and $w^\star_{i,\nu}$). We naturally assume that $v^\star_{\bullet,\bullet}$ $w^\star_{\bullet,\bullet}$ are mutually independent and independent of $\Xi_0, u_\bullet, \mathcal{V}_\bullet, \mathcal{W}_\bullet$.

Let $\theta > 0$ be a real constant and let $i \in \mathbb{N}$. Denote

$$\theta_i \triangleq (\theta \vee \tau_{i-1}) \wedge \tau_i,$$

$$\hat{x}_i \triangleq \begin{cases} x_i & \text{if } \theta_i = \tau_i \\ \infty & \text{otherwise} \end{cases}$$

and

$$\tilde{x}_i \triangleq \begin{cases} \infty & \text{if } \theta_i = \tau_i \\ x_i & \text{otherwise.} \end{cases}$$

Note that $\bar{\xi}_{\theta_i}$ is uniquely determined by $\bar{\xi}_{\tau_{i-1}}$, $\hat{x}$ and θ.

Since $(\bar{\xi}_t)_{t\geq 0}$ is of a pure-jump type and (trivially) Markov, it is (by [3], Theorem 12.14.) strongly Markov; hence we may (by [3] Proposition 6.13.) assume the existence of variables $\hat{u}_i \sim \mathrm{U}(0,1), \tilde{u}_i \sim \mathrm{U}(0,1)$ such that

$$\hat{u}_i \perp\!\!\!\perp \bar{\xi}_{\tau_{i-1}}, \qquad \hat{x}_i = \hat{f}(\bar{\xi}_{\tau_{i-1}}, \hat{u}_i) \tag{18}$$

$$\tilde{u}_i \perp\!\!\!\perp \bar{\xi}_{\theta_i}, \hat{u}_i, \qquad \tilde{x} = \tilde{f}(\bar{\xi}_{\theta_i}, \tilde{u}) \tag{19}$$

for some measurable functions $\hat{f}, \tilde{f}$.

It can be easily verified that the distribution of Ξ will remain unchanged if we assume that x_i is defined by (18) and (19) (instead of its original definition by means of u_i, assumed in Section A.3), provided that $\hat{u}, \tilde{u}$ are independent of the underlying variables (12).[17] Moreover, it will be

$$\tilde{u} \perp\!\!\!\perp \bar{\Xi}_{\theta_i} \tag{20}$$

given our assumption (to see it, note that $\bar{\Xi}_{\theta_i}$ is a function of $\bar{\xi}_{\theta_i}$ and the variables (12)).

Further, note that, by our definition of the dynamics and our reconstruction,

$$\Xi|_{(\theta_i,\infty)} = \Upsilon_i(\dot{\Xi}_{\theta_i}, U_i), \qquad U_i \triangleq (\bar{U}_i, \tilde{U}_i, \tilde{u}_i)$$

for some mapping Υ_i, where

$$\bar{U}_i \triangleq (u_{i+1}, \mathcal{V}_{i+1}, \mathcal{W}_{i+1}, v_{i+1}, w_{i+1}, u_{i+2}, \mathcal{V}_{i+2}, \dots),$$

$$\tilde{U}_i \triangleq (\tilde{\mathcal{V}}, \tilde{\mathcal{W}}, \tilde{v}^\star_\bullet, \tilde{w}^\star_\bullet),$$
$$\tilde{\mathcal{V}} \triangleq \mathcal{V}_i|_{(\vartheta,\infty)} \ominus_1 \vartheta, \qquad \vartheta \triangleq \tau_i - \theta_i,$$
$$\tilde{v}^\star_\nu \triangleq v^\star_{\kappa_\nu}|_{(s_\nu,\infty)} \ominus s_\nu, \qquad s_\nu \triangleq \vartheta \rho_i(a_{\theta_i,\nu}),$$

and where κ_ν is the ordering of $a_{\vartheta,\nu}$ in $A_{\tau_{i-1}}$ ($\tilde{\mathcal{W}}$ and $w^\star$ are defined symmetrically).

Let us examine $\mathcal{L}\left(U_i \middle| \bar{\Xi}_{\theta_i}\right)$ now: Note first that, thanks to the independence of variables (12),

$$\mathcal{L}\left(\bar{U}_i \middle| \bar{\Xi}_{\theta_i}, \tilde{u}_i, \tilde{U}_i\right) = \mathcal{L}\left(\bar{U}_i\right) \tag{21}$$

where $\mathcal{L}\left(\bar{U}_i\right)$ is independent of i. Take $i \in \mathbb{N}$ and $k = (k_2, k_3, \dots)$, $l = (l_2, l_3, \dots)$, $k_\nu, l_\nu \in \mathbb{N}$, $\nu > 1$ and define the set

$$S_{i,k,l} = [\kappa_\bullet = k_\bullet, \lambda_\bullet = l_\bullet].$$

If $(\heartsuit_i)$ held true and, in addition, $\tilde{u}_i$ was deterministic, then $\theta_i, \sigma_i, \rho_i$ would also be deterministic, so it would be easy to see that $\tilde{\mathcal{V}}_i$ would be a P.p.p. with the same intensity as $\mathcal{V}_1$ and that $(\tilde{v}^\star_\nu)_{\nu>1}$ would be a collection of independent unit P.p.p.'s on $\mathbb{R}_+$, (symmetrically $\tilde{\mathcal{W}}_i$, $\tilde{w}^\star$); moreover, all the components of $\tilde{U}_i$ would be mutually independent. Summarised,

$$\mathcal{L}\left(\tilde{U}_i \middle| \bar{\Xi}_{\theta_i}, \tilde{u}_i\right) = \mathfrak{d} \tag{22}$$

on $S_{i,j,k}$ for some distribution $\mathfrak{d}$ depending neither on i nor on $S_{i,j,k}$. Hence, by Lemma A.2 (and thanks to the facts that $S_{i,j,k}$ is measurable w.r.t. $\sigma(\bar{\Xi}_{\theta_i}, \tilde{x}_i)$ and that sets $S_{i,j,k}$ cover Ω), relation (22) holds on the whole Ω.

Now we are able to apply the Chain Rule ([3], Proposition 6.8.) to (21), (22) and (20) to prove that

$$\mathcal{L}\left(U_i \middle| \bar{\Xi}_{\theta_i}\right) = \mathfrak{D} \tag{23}$$

[17] Note that x_i is uniquely determined by $\hat{x}_i$ and $\tilde{x}_i$.

for a distribution $\mathfrak{D}$ not depending on i.

Finally, denote

$$\iota \triangleq \min\{i : \theta \le \tau_i\},$$
$$U \triangleq U_\iota, \qquad \Upsilon \triangleq \Upsilon_\iota.$$

Since $\theta_\iota = \theta$, we obtain that

$$\dot{\Xi}|_{(\theta,\infty)} = \Upsilon(\dot{\Xi}_\theta, U). \tag{24}$$

Moreover, it follows from (23) and from Lemma A.2 (applied to $S_i \triangleq [\iota = i]$, $i \in \mathbb{N}$) that

$$\mathcal{L}\left(U \middle| \bar{\Xi}_\theta\right) = \mathfrak{D}.$$

Since the distribution $\mathfrak{D}$ does not depend on $\bar{\Xi}_\theta$, necessarily $\mathfrak{D} = \mathcal{L}(U)$, i. e., U and $\bar{\Xi}_\theta$ are independent which is, together with (24), sufficient for the validity of (iii) by the equivalent definition of Markov property given by [3] (see the start of Chapter 8) and by Lemma 1.12. from the same book. □

Proof of Proposition 4.1. Our aim is to determine the distribution of the components of A_θ given Ξ_0 and $\bar{\xi}_\theta$; however, we (temporarily) add the opposite order book B_θ among the conditioning elements. We will later find this addition useful when proving the conditional independence of the books.

Let $1 \le i \le \tilde{I} + 1$. The fact that the prices of all the limit orders waiting at the time $\tilde{\tau}_i$ must either equal $\tilde{a}_i$ or be not smaller than $\tilde{a}_{i-1}$,[18] implies that $\dot{A}_i$ is uniquely determined by $\tilde{a}_i$ and the set of the orders with limit prices no less than $\tilde{a}_{i-1}$ waiting at time $\Delta\tilde{\tau}_i$, which is itself determined by

$$\dot{A}_{i-1}, \Delta\tilde{\tau}_i, \tilde{\mathcal{Y}}_i \text{ and } \tilde{v}_i$$

where

$$\tilde{\mathcal{Y}}_i \triangleq \mathcal{Y}_i|_{\mathbb{R}_+ \times (\tilde{a}_{i-1} - a_{\tau_{i-1}}, \infty) \times \mathbb{R}_+}$$

and

$$\tilde{v}_i = \tilde{v}_i^\theta \triangleq (v_{i,k_i}, v_{i,k_i+1}, \dots), \qquad k_i \triangleq 1 + \max\{\nu : a_{\tau_{i-1},\nu} \le \tilde{a}_{i-1}\}.$$

Applying this procedure repeatedly and putting $\tilde{a}_\nu - a_{\tau_\nu} = 0$ and $k_\nu \triangleq 2$ for $\nu > \tilde{I} + 1$, we find that A_θ (which equals to $\dot{A}_{\tilde{I}+1}$) is a function of $\dot{A}_0$, $\bar{\xi}_\theta$, θ and $\tilde{\mathcal{Y}}_1, \tilde{v}_1, \tilde{\mathcal{Y}}_2, \tilde{v}_2, \dots$. Since, by our construction,

$$\tilde{\mathcal{Y}}_i = G(\mathcal{V}_i|_{(0,\Delta\tau_i)\times(\tilde{v}_i,\dot{v}_i)\times(0,1)}; \bar{\xi}_{\tau_i})$$

where $\tilde{v}_i \triangleq \Phi_i(\tilde{a}_{i-1} - a_{\tau_{i-1}})$, we are getting that $\tilde{\mathcal{Y}}_i$ is a function of $\bar{\xi}_{\tau_i}, \Delta\tau_i$ and

$$\tilde{\mathcal{V}}_i = \tilde{\mathcal{V}}_i^\theta \triangleq \mathcal{V}_i|_{\mathbb{R}_+ \times (\tilde{v}_i, \infty) \times \mathbb{R}_+} \ominus_2 \tilde{v}_i$$

hence

$$A_\theta \text{ is a function of } \dot{A}_0,\ \bar{\xi}_\theta,\ \theta \text{ and } \tilde{\mathcal{V}}_1, \tilde{v}_1, \tilde{\mathcal{V}}_2, \tilde{v}_2, \dots.$$

The following Auxiliary Assertion proves a useful – quite intuitive – result:

[18] Indeed, if $\tilde{a}_{i-1} = \tilde{a}_i$ then the assertion is trivial, if $\tilde{a}_{i-1} > \tilde{a}_i$ then necessarily $a_{\tilde{\tau}_i^-} = \tilde{a}_{i-1}$ so any order between $\tilde{a}_i$ and $\tilde{a}_{i-1}$ waiting at $\tilde{\tau}_i$ would have to arrive at $\tilde{\tau}_i$ which is impossible by our assumption following the Proposition 3.1.

Auxiliary Assertion. For each $i \in \mathbb{N}$, $\tilde{\mathcal{V}}_i^\theta$ is a P.p.p. with unit intensity and $\tilde{v}_i^\theta$ is a vector of i.i.d. unit exponential variables. Moreover,

$$\tilde{\mathcal{V}}_1^\theta, \tilde{v}_1^\theta, \tilde{\mathcal{V}}_2^\theta, \tilde{v}_2^\theta, \dots$$

are mutually independent and independent of $\bar{\xi}_\theta, \Xi_0, B_\theta, \theta$.

Proof of A.A. First we prove, by induction, that the A.A. holds given that

$$\tilde{\tau}_{k-1} \leq \theta \leq \tilde{\tau}_k \tag{25}$$

for some $k \in \mathbb{N}$.

Assume first that $k = 1$. In that case, $\tilde{I} = 0$; hence $\tilde{\mathcal{V}}_i = \mathcal{V}$ and $\tilde{v}_i = v_i$ for each i. Since, in addition, θ is a function of (u_1, ξ_0) and B_θ is a function of Ξ_0, $\mathcal{W}_\bullet$ and $w_\bullet$, the elements $\tilde{\mathcal{V}}_\bullet$, $\tilde{v}_\bullet$ are independent of $\bar{\xi}_\theta, \Xi_0, B_\theta, \theta$ hence the A.A. holds for $k = 1$.

Let now $k > 1$ and assume that the A.A. holds for $k - 1$. For better readability, let us agree to write $\tilde{a}'$ instead $\tilde{a}^{\tau_{k-1}}$, $\tilde{\mathcal{V}}_i'$ instead of $\tilde{\mathcal{V}}_i^{\tau_{k-1}}$ etc., and to omit index θ at $\tilde{a}^\theta$, $\tilde{\mathcal{V}}_i^\theta$ etc. (i. e., $\tilde{a}$ will stand for $\tilde{a}^\theta$ etc.).

By the induction hypothesis with $\theta = \tau_{k-1}$, we have

$$\tilde{\mathcal{V}}_\bullet', \tilde{v}_\bullet' | \Xi_0, \bar{\xi}_{\tau_{k-1}}, B_{\tau_{k-1}} \sim \mathfrak{P} \otimes \mathfrak{E}$$

where $\mathfrak{P}$ is a distribution of a P.p.p. on $\mathbb{R}_+ \times \mathbb{R}_+ \times (0,1)$ with unit intensity and $\mathfrak{E}$ is a distribution of a sequence, indexed by $\mathbb{N}$, of independent unit exponential variables, which implies

$$\tilde{\mathcal{V}}_\bullet', \tilde{v}_\bullet' \perp\!\!\!\perp \Xi_0, \bar{\xi}_{\tau_{k-1}}, B_{\tau_{k-1}}.$$

Further, since $\tilde{\mathcal{V}}_\bullet', \tilde{v}_\bullet'$, viewed as functions of the underlying elements (12), are constant in $u_k, \mathcal{W}_k$ and w_k, it also holds that

$$\tilde{\mathcal{V}}_\bullet', \tilde{v}_\bullet' \perp\!\!\!\perp u_k, \mathcal{W}_k, w_k;$$

with respect to our construction (Section A.3) we get

$$u_k, \mathcal{W}_k, w_k \perp\!\!\!\perp \Xi_0, \bar{\xi}_{\tau_{k-1}}, B_{\tau_{k-1}}$$

we have a mutual independence of $(u_k, \mathcal{W}_k, w_k)$, $(\Xi_0, \bar{\xi}_{\tau_{k-1}}, B_{\tau_{k-1}})$ and $(\tilde{\mathcal{V}}_\bullet', \tilde{v}_\bullet')$ implying

$$\tilde{\mathcal{V}}_\bullet', \tilde{v}_\bullet' \perp\!\!\!\perp \Upsilon, \qquad \Upsilon \triangleq (\Xi_0, \bar{\xi}_{\tau_{k-1}}, B_{\tau_{k-1}}.u_k, \mathcal{W}_k, w_k).$$

Assume now, for a while, that

($\sharp$) $\Upsilon, \tilde{v}_\bullet'$ are deterministic

and let us investigate the distribution of $\tilde{\mathcal{V}}_\bullet'$.

We begin with the case of $\chi_k = \mathfrak{a}^+$. It follows from the definition of the dynamics that, in this case,

$$b_{\tau_k} = b_{\tau_{k-1}}, \qquad a_{\tau_k} = \aleph \tag{26}$$

where $\aleph$ is the first limit order except of the ask waiting at τ_k^-, i. e.,

$$\aleph = \min\{p : p \in \dot{A}'_0 \cup \breve{A}', p \text{ is not cancelled}\} \\ \wedge \min_i \min\{p : p \in R_i, p \text{ is not cancelled}\}$$

where R_i is a p.p. formed by the second coordinates of $(a_{\tau_{i-1}} \oplus Y_i)|_{\mathbb{R}_+ \times (\tilde{a}'_{i-1}, \infty)}$. Denoting $(t_{i,\nu}, \pi_{i,\nu}, \eta_{i,\nu})_{\nu \in \mathbb{N}}$ the points of $\tilde{\mathcal{V}}'_i|_{(0.\Delta\tau_i) \times \mathbb{R}_+ \times (0,1)}$, $i \leq k$, we are getting, by our reconstruction, that

$$R_i \stackrel{\triangle}{=} (a_{\tau_{i-1}} + \Phi_i^{[-1]}(\Phi_i(\tilde{a}'_{i-1} - a_{\tau_{i-1}}) + \pi_\nu))_{\nu \in \mathbb{N}}. \tag{27}$$

In the following text, let us agree to view R_i as a marked point process with marks $(t_\nu, \eta_\nu)_{\nu \in \mathbb{N}}$. Since, under ($\sharp$), each R_i is (by Lemma A.3 (ii)) a P.p.p. on $(\tilde{a}'_{i-1}, \infty)$ with intensity $\Phi_i \oplus a_{\tau_{i-1}}$ and with i.i.d. marks (note that $\tilde{a}_\bullet$ are deterministic under ($\sharp$)) and since $\tilde{\mathcal{V}}'_1, \dots, \tilde{\mathcal{V}}'_k$ are independent (by the induction hypothesis), we have that

$$\underline{R}_1^p, \dots, \underline{R}_k^p, \perp\!\!\!\perp \overline{R}_1^p, \dots, \overline{R}_k^p,$$

$$\underline{R}_i^p \stackrel{\triangle}{=} R_i|_{[0,p)} \qquad \overline{R}_i^p \stackrel{\triangle}{=} R_i|_{[p,\infty)},$$

for any $p \in \mathbb{R}$, with the consequence that

$$\underline{R}_1^p, \dots, \underline{R}_k^p \perp\!\!\!\perp S_1^p, \dots, S_k^p \tag{28}$$

for any $p \in \mathbb{R}$, where

$$S_i^p \stackrel{\triangle}{=} \{(t, \Phi_i(q-p), \eta) : q \in \overline{R}_i,\ (t,\eta) \text{ is a mark of } q\}.$$

Note that $S_1^p, \dots, S_k^p$ are mutually independent and, by Lemma A.3 (ii),

$$S_i^p \sim \mathfrak{p}, \qquad 1 \leq i \leq k, \tag{29}$$

where $\mathfrak{p}$ is the distribution of a unit P.p.p. on $(0, \Delta\tilde{\tau}_i) \times \mathbb{R}_+ \times (0,1)$.

Now introduce a m.p.p. $R \stackrel{\triangle}{=} R_1 \cup \dots \cup R_k$ with an additional mark determining an index of a "source" process of each point (i. e., the mark of a point $\pi \in R$ equals i iff $\pi \in R_i$). Let r_p be a stochastic process (with "time" p) whose jump times coincide with points of R and the magnitudes of the jumps are equal to the corresponding uniquely coded marks (the existence of such coding is guaranteed by [3], Lemma 1.12). Since, r is a process with independent increments by (27), it is clearly Markov. Moreover, since r is a pure-jump type process, it is even strong Markov (by [3], Theorem 12.14.); therefore and because $\aleph$ is a $\sigma(r_p)$-optional time (indeed, to determine whether or not $\aleph \leq p$, only the history of r up to p and deterministic variables are needed), relations (28) and (29) hold with $p = \aleph$. Further, since optional times are measurable with respect to "their" sigma fields (see [3], Lemma 7.1), we are getting that $S_1^\aleph, \dots, S_k^\aleph | \aleph \sim \mathfrak{p}^k$ implying

$$S_1^\aleph, \dots, S_k^\aleph \perp\!\!\!\perp \aleph. \tag{30}$$

Now, for any $i \leq k$ and $p \in \mathbb{R}$, put

$$T_i^p \triangleq \tilde{\mathcal{V}}_i'|_{[\Delta\tilde{\tau}_i,\infty)\times(\Phi_i(p-\tilde{a}'_{k-1}),\infty)} \ominus_2 \Phi_i(p - \tilde{a}'_{k-1}),$$

(i. e., the unused part of $\tilde{\mathcal{V}}_i'$) and note that

$$\tilde{\mathcal{V}}_i = \begin{cases} S_i^{\aleph} \cup T_i^{\aleph} & i \leq k \\ \tilde{\mathcal{V}}_i' & i > k. \end{cases}$$

Since, by assertion (ii) of Lemma A.3 (with $\aleph$ made temporarily deterministic),

$$T_1^{\aleph}, \dots, T_k^{\aleph} | S_1^{\aleph}, \dots, S_k^{\aleph}, \aleph \tag{31}$$

are independent unit Poisson, we are getting, by the Chain Rule ([3], Proposition 6.8.) applied to (30), (31) and to the relation $(\tilde{\mathcal{V}}_i)_{i>k} \perp\!\!\!\perp (\tilde{\mathcal{V}}_i)_{i\leq k}, \pi$ (following from the definitions) that

$$\tilde{\mathcal{V}}_\bullet | \xi_{\tau_k} = \mathfrak{P} \tag{32}$$

given the assumption ($\sharp$) and the case $\chi_k = \mathfrak{a}^+$ (note that ξ_{τ_k} is a function of $\aleph$ and deterministic elements).

If, on the other hand, $\chi_k \neq \mathfrak{a}^+$, the situation is much simpler: in this case, $\tilde{\mathcal{V}}_\bullet = \tilde{\mathcal{V}}_\bullet'$ and ξ_{τ_k} is a function of deterministic elements $\bar{\xi}_{\tau_{k-1}}$, x_i, $B_{\tau_{k-1}}$, $\mathcal{W}_k$ and w_k, hence ξ_{τ_k} is deterministic implying (32) given $\chi_k \neq \mathfrak{a}^+$, too.

Finally, releasing ($\sharp$) and assuming the true distribution of Υ, we are getting, by assertion (ii) of Lemma A.1, that

$$\tilde{\mathcal{V}}_\bullet | \xi_{\tau_k}, \Upsilon, \tilde{v}' \sim \mathfrak{P}$$

implying

$$\tilde{\mathcal{V}}_\bullet | \xi_{\tau_k}, \Upsilon \sim \mathfrak{P} \tag{33}$$

because $\mathfrak{P}$ does not depend on $\tilde{v}'$.

Now proceed to the distribution of $\tilde{v}_\bullet$: let us make another temporary assumption that

($\flat$) $\Upsilon, \tilde{\mathcal{V}}_\bullet'$ are deterministic

and, for any $\kappa_1, \dots, \kappa_k \in \mathbb{N}$, define set (of elementary events)

$$\mathcal{S}_{\kappa_1,\dots,\kappa_k} \triangleq [k_1 = \kappa_1, \dots, k_k = \kappa_k].$$

On each $\mathcal{S}_{\kappa_1,\dots,\kappa_k}$, clearly

$$\tilde{v}_i = \{v_{i,\kappa_i}, v_{i,\kappa_i+1}, \dots\}, \qquad 1 \leq i \leq k,$$

and

$$\xi_{\tau_k} = h(v_{1,k_1'}, \dots, v_{1,k_1}, v_{2,k_2'}, \dots, v_{2,k_2}, v_{k,k_2'}, \dots, v_{k,k_2})$$

for some deterministic function h. Hence, by the independence of components of $v_\bullet$ and thanks to the fact that $k_i = 1$ for $i > k$,

$$\tilde{v}_\bullet | \xi_{\tau_k} \sim \mathfrak{E} \tag{34}$$

on $\mathcal{S}_{\kappa_1,\dots,\kappa_k}$ by Lemma A.2. Since the sets $\mathcal{S}_\bullet$ cover the entire probability space Ω, (34) holds universally. Further, after releasing ($\flat$), we are getting (by Lemma A.1 (ii), that

$$\tilde{v}_\bullet | \xi_{\tau_k}, \Upsilon, \tilde{\mathcal{V}}'_\bullet \sim \mathfrak{E} \tag{35}$$

which yields

$$\tilde{v}_\bullet | \xi_{\tau_k}, \Upsilon, \tilde{\mathcal{V}}_\bullet \sim \mathfrak{E} \tag{36}$$

because $\tilde{\mathcal{V}}_\bullet$ is a function of the elements in the condition of (35).

Now, by an application of the Chain rule ([3], Proposition 6.8) to (36) and (33), we finally get

$$\tilde{\mathcal{V}}_\bullet, \tilde{v}_\bullet | \xi_{\tau_k}, \Upsilon \sim \mathfrak{P} \otimes \mathfrak{E}.$$

Moreover, since each θ fulfilling (25) is a function of $\bar{\xi}_{\tau_k}$ which itself is a function of Υ and ξ_{τ_k} and since B_θ is a function of $B_{\tau_{k-1}}$, $\mathcal{W}_k$ and w_k, we get

$$\tilde{\mathcal{V}}_\bullet, \tilde{v}_\bullet | \Xi_0, \theta, \bar{\xi}_\theta, B_\theta \sim \mathfrak{P} \otimes \mathfrak{E},$$

i. e., the assertion of the A.A. in the case that θ fulfils (25).

Finally, let θ be a general optional time. For each $k \in \mathbb{N}$, define

$$\mathcal{S}_k \triangleq [\theta \in [\tau_{i-k}, \tau_k)]. \tag{37}$$

Since, on each $\mathcal{S}_k$,

$$\theta = \theta_k, \qquad \theta_k \triangleq (\theta \vee \tau_{k-1}) \wedge \tau_k$$

and since θ_k fulfils (25), we are getting (37) on each S_k by Lemma A.2. Since $\mathcal{S}_\bullet$ cover Ω, the Auxiliary Assertion is proven. □

Now, in line with the discussion succeeding Lemma A.1, assume, for some time, that

($\diamondsuit$) $\bar{\xi}_\theta, \Xi_0, B_\theta$ and θ are deterministic while $\tilde{\mathcal{V}}_1, \tilde{v}_1, \tilde{\mathcal{V}}_2, \tilde{v}_2$ keep their distribution (see the A.A.)

and try to derive the conditional distribution of A_θ given $\Xi_0, \theta, \bar{\xi}_\theta, B_\theta$; before starting, note that all the elements in the condition are constant under ($\diamondsuit$) hence we are looking for an ordinary (unconditional) distribution. We shall proceed gradually according to our decomposition of A_θ.

The easiest work is with a_θ: since it may be determined directly from $\bar{\xi}_\theta$, it is deterministic given ($\diamondsuit$) i. e., *(-)* is proven.

Let us proceed to the sets $L_\bullet$: Fix $1 \le i \le \tilde{I} + 1$ and denote by K_i the p.p. of the *relative* limit prices of the orders originated in Λ_i having survived until τ_i, i. e.,

$$K_i = \{p : (t, p, c) \in \mathcal{Y}_i : p > \tilde{a}_{i-1} - a_{\tau_{i-1}}, c + t > \Delta\tilde{\tau}_i\}.$$

Viewing the second coordinates of $\mathcal{Y}_i$ as its ground process and regarding the pair of the first and third coordinates as marks (note that then the first coordinates are uniformly distributed), we get by Lemma A.3 (iii) that

$$\mathcal{Y}_i \stackrel{d}{=} \gamma(\mathcal{Y}'_i), \qquad \gamma(t,p,u) \triangleq (t,p,-\ln(u)/\rho_i(p))$$

where $\mathcal{Y}'_i$ is a P.p.p. with intensity

$$\mu' \triangleq \ell \otimes \mathfrak{F}_i \otimes \ell$$

on $(0,\Delta\tau_i)\times\mathbb{R}_+\times(0,1)$ implying that

$$K_i \stackrel{d}{=} K'_i, \qquad K'_i = \{p : (t,p,u)\in\mathcal{Y}'_i : p > \tilde{a}_{i-1} - a_{\tau_{i-1}}, u < \exp\{-\rho_i(p)(\Delta\tilde{\tau}_i - t)\},$$

which further yields that, for any $S \in \mathfrak{B}(\mathbb{R})$,

$$K_i S = \mathcal{Y}'_i S', \qquad S' \triangleq \{(t,p,u) : p \in S, p > \tilde{a}_{i-1} - a_{\tau_{i-1}}, u < \exp\{-\rho_i(p)(\Delta\tilde{\tau}_i - t)\}.$$

Therefore and since $\mathcal{Y}'_i$ is a P.p.p.,

$$K_i S \sim \text{Poisson}\left(\mu'(S')\right)$$

and we get, by a textbook integration, that

$$\begin{aligned}\mu'(S') &= \int_{S\cap(\tilde{a}_{i-1}-a_{\tau_{i-1}},\infty)} \int_0^{\Delta\tilde{\tau}_i} \int_0^{\exp\{-\rho_i(p)(\Delta\tilde{\tau}_i - t)\}} \phi_i(p)\,\mathrm{d}u\,\mathrm{d}t\,\mathrm{d}p \\ &= \int_{S\cap(\tilde{a}_{i-1}-a_{\tau_{i-1}},\infty)} \int_0^{\Delta\tilde{\tau}_i} \exp\{-\rho_i(p)(\Delta\tilde{\tau}_i - t)\}\phi_i(p)\,\mathrm{d}p\,\mathrm{d}t \\ &= \int_S \mathbf{1}_{(\tilde{a}_{i-1}-a_{\tau_{i-1}},\infty)}(p) f_i(p)\,\mathrm{d}p.\end{aligned}$$

Hence, by the uniqueness criterion for simple point processes [3, Theorem 12.8], K_i is a P.p.p. with density

$$\mathbf{1}_{(\tilde{a}_{i-1}-a_{\tau_{i-1}},\infty)}(p) f_i(p)$$

and, moreover,

$$K_i \perp\!\!\!\perp K_1,\dots,K_{i-1},e_1,\dots,e_i \tag{38}$$

(which is true because the variables on the r.h.s. depend on underlying variables, independent of those by means of which K_i is defined).

Further, for any $j\in\mathbb{N}$, $i < j \le \tilde{I}+1$, denote by $P_i^j = [p_{i,j,1}, p_{i,j,2},\dots]$ the p.p. of the *absolute* limit prices of the orders from K_i uncancelled at $\tilde{\tau}_j$ and note that $P_i^i = K_i \oplus a_{\tau_{i-1}}$. Let, for each $\nu\in\mathbb{N}$, $k_{i,j,\nu}$ be ordering of $p_{i,j,\nu}$ in $A_{\tau_{j-1}}$ and note that it follows from our (re)definitions that an order (with the limit price) $p_{i,j,\nu}$ is uncancelled at $\tilde{\tau}_j$ if

$$e_{j,k_{i,j,\nu}}/\rho_i(a_{\tau_{i-1},k_{i,j,\nu}} - a_{\tau_{i-1}}) \ge \Delta\tilde{\tau}_j.$$

Taking $e_{j,k_{i,j,\nu}}$ as a mark of $p_{i,j,\nu}$ for each ν, it follows from Lemma A.3 (iii) that, whenever P_i^{j-1} is a P.p.p. with a density h_{j-1}, then, for each $S \in \mathfrak{B}(\mathbb{R})$,

$$P_i^j S = \text{Poisson}\left(\int_S \int_0^{\exp\{-\rho_j(p-a_{\tau_{j-1}})\Delta\tilde{\tau}_j\}} h_{j-1}(p)\,\mathrm{d}u\,\mathrm{d}p\right)$$
$$= \text{Poisson}\left(\int_S h_{j-1}(p)\exp\{-\rho_j(p-a_{\tau_{j-1}})\Delta\tilde{\tau}_j\}\,\mathrm{d}p\right)$$

i. e., P_i^j is a P.p.p. with density $h_{j-1}(p)\exp\{-\rho_j(p-a_{\tau_{j-1}})\Delta\tilde{\tau}_j\}$. Applying this procedure to $j = i+1, \ldots, \tilde{I}+1$ and noting that $L_i = P_i^{\tilde{I}+1}$ we finally get that L_i is a P.p.p. with a density

$$\mathbf{1}_{(\tilde{a}_{i-1},\infty)}(p) f_i(p - a_{\tau_{i-1}})\epsilon(i,p). \tag{39}$$

We further show that, for any $1 \le j \le \tilde{I}+1$,

$$P_0^j, \ldots, P_j^j \text{ are mutually independent.} \tag{40}$$

Indeed: assume that (40) holds for $j-1$, i. e.,

$$P_0^{j-1}, \ldots P_{j-1}^{j-1} \tag{41}$$

are independent. If the variables (41) were deterministic then $k_{i,j,\bullet}$, $i < j$, would be deterministic, too, hence the vectors $e_i^j \stackrel{\triangle}{=} e_{j,k_{i,j,\bullet}}$, $i < j$, would be mutually independent i.i.d. vectors of unit exponential variables. Therefore, by Lemma A.1 (ii), $(e_i^j)_{j<i}$ are mutually independent and independent of variables (41). The independence (40) now follows from the fact that each P_i^j, $i < j$, is a function of (P_i^{j-1}, e_i^j) and from the independence of P_j^j of variables (41) and e_j.

Since $L_i = P_i^{\tilde{I}+1}$, we have proven the independence of $L_1, \ldots, L_{\tilde{I}+1}$ further implying that L is Poisson with density equal to the sum of the densities (39) which proves *(i)*.

Let us proceed to D: let $1 \le i \le \tilde{I}+1$ and denote by $D_i = \left[\, d_1^i, d_2^i, \ldots\right]$ the set of all the orders contained in $\dot{A}_0$ and uncancelled at $\tilde{\tau}_i$; it follows from the independence of e_i and $\dot{A}_{i-1}, D_{i-1}$ that D_i is a heterogeneous thinning of D_{i-1} with parameters

$$\exp\{-\rho_i(d_1^{i-1} - a_{\tau_{i-1}})\Delta\tilde{\tau}_i\}, \exp\{-\rho_i(d_2^{i-1} - a_{\tau_{i-1}})\Delta\tilde{\tau}_i\}, \ldots$$ [19]

which, used $(\tilde{I}+1)$-times, gives that $D = D_{\tilde{I}+1}$ is an h-thinning of $\dot{A}_0$ with parameters

$$\epsilon(0, \dot{a}_1), \epsilon(0, \dot{a}_2), \ldots$$

where $\dot{a}_1 < \dot{a}_2 < \ldots$ are the points of $\dot{A}_0$, which is nothing else but *(ii)*, Moreover, it could be shown similarly to the proof of (40) that $D \perp\!\!\!\perp L_1, \ldots, L_{\tilde{I}+1}$ hence $D \perp\!\!\!\perp L$.

Let us examine E now: it follows from the definition of E that for p to be its point, two conditions have to be satisfied:

[19] To see it, view $\dot{A}_0$ and D_{i-1} as deterministic, note that the components of e_i corresponding to the points of D_{i-1} form a vector of i.i.d. unit exponential variables; i. e., it is easy to check that D_i formally fulfils the condition defining an h-thinning, and use Lemma A.1 (ii).

1. It has to be $p = \breve{a}_\nu$ for some $\nu \leq \tilde{I}$.

2. It must have survived since $\tilde{\tau}_{\gamma_\nu - 1}$ until θ.

The distribution of E (described by *(iii)*) and its independence of L, D may be determined similarly to our proofs of the analogous properties of D.

Having derived the distribution of L, D and E and having proven their independence – i. e., *(iv)* – given ($\diamondsuit$), we have fully described the distribution of A_θ given ($\diamondsuit$).

Now we may stop assuming ($\diamondsuit$). Due to the independence proven by the Auxiliary Assertion, we may use Lemma A.1 (ii) to get that the formulas describing the distribution of A_θ given $\Xi_0, \theta, \bar{\xi}_\theta, B_\theta$ remain valid even without ($\diamondsuit$) (note that the independence of L, D and E given ($\diamondsuit$) corresponds to their conditional independence given $\Xi_0, \theta, \bar{\xi}_\theta, B_\theta$). Moreover, since $\mathcal{L}\left(A_\theta \,\middle|\, \Xi_0, \theta, \bar{\xi}_\theta, B_\theta\right)$ does not depend on B_θ, it is simultaneously $\mathcal{L}\left(A_\theta \,\middle|\, \Xi_0, \theta, \bar{\xi}_\theta\right)$.

The validity of (i), (iii), (iv) with C' and (ii') could be proven analogously by assuming ($\diamondsuit$) with B_0 instead of Ξ_0 (the distribution of $D|C'$ might be derived analogously as the one of L_i for some i). The point ($*$) would follow by a mirroring according to the price, ($\triangle$) follows from the fact that the conditional distribution of A_θ given both C and C' equals to the one given (C, B_θ), (C', B_θ), respectively. □

P r o o f o f P r o p o s i t i o n 4.2. Since, by definition of the dynamics, $\Delta a_{\tau_k} = y_k$ if $\chi_k = \mathfrak{a}^-$ and $\Delta a_{\tau_k} = 0$ if $\chi_k \in \{\mathfrak{b}^+, \mathfrak{b}^-\}$, the relation (4) and the second branch of (5) are trivial. It remains to derive the distribution of Δa_{τ_k} given $\chi_k = \mathfrak{a}^+$ in which case

$$\Delta a_{\tau_k} = y'_k$$

where y'_k is the distance of $a_{\tau_{k-1}}$ to the next point of $A_{\tau_k^-}$ (if there is any) or equals to ∞ (if $|A_{\tau_k^-}| = 0$). Hence, we first have to determine $\mathcal{L}\left(A_{\tau_k^-} \,|C_k\right)$, which task is, however, very similar to deriving $\mathcal{L}\left(A_{\tau_k} | \dots\right)$ in the proof of Proposition 4.1 with the "only" difference that now we have to get rid of the (possible) jump of a at τ_k (which is actually the reason why $\tilde{a}^{\tau_{k-1}}$, $\breve{A}^{\tau_{k-1}}$, etc., appear in the formulas); hence, we will not go through all the details but we only note that the proof would also be done by means of a decomposition of $A_{\tau_k^-}$ into

- the ask,
- a conditionally Poisson p.p. L' with the density $\grave{F}$,
- an h-thinning D' of $\dot{A}_0^{\tau_{k-1}}$ with parameters $\epsilon^{\tau_k}(0, \bullet)$ and
- an h-thinning E' of $\breve{A}^{\tau_{k-1}}$ with parameters $\epsilon^{\tau_k}(\bullet, \bullet)$,

all the four elements being conditionally independent. Clearly,

$$y'_k > p \Leftrightarrow (A_{\tau_k^-})(a_{\tau_k^-}, a_{\tau_k^-} + p] = 0.$$

Since, for any interval I,

$$\mathbb{P}[L'I = 0|C_k] = \exp\{-\eta I\}$$

where η is the (conditional) intensity measure of L' (whose density is $\grave{F}$, as we have already proven) and since, for any h-thinning Y of a p.p. $X = [x_1, x_2, \dots]$ with parameters $p_1, p_2, \dots,$

$$\mathbb{P}[YI = 0|X] = \prod_{\nu \in \mathbb{N}, x_\nu \in I} (1 - p_\nu)$$

we are getting (i) (also using the conditional independence of L', D' and E').

The proof of (i') is similar (see also the proof of Proposition 4.1). The part $(*)$ may be derived by mirroring. The conditional independence in $(\triangle)$ trivially follows from the fact that always at least one of the values a_{τ_k} and b_{τ_k} is conditionally constant given $\xi_{\tau_{k-1}}, \chi_k$. $\square$

ACKNOWLEDGEMENT

This work was supported by Grants No. 402/06/1417, P402/10/0956 and P402/10/1610 of the Czech Science Foundation and by Project No. LC06075 of the Ministry of Education, Youth and Sports. The author is grateful to both of the anonymous referees for valuable comments and corrections.

(Received January 25, 2011)

REFERENCES

[1] D. J. Daley and D. Vere-Jones: An Introduction to the Theory of Point Processes. Second edition. Springer, New York 2003.

[2] J. Hoffmann-Jørgenson: Probability with a View Towards to Statistics I. Chapman and Hall, New York, 1994.

[3] O. Kallenberg: Foundations of Modern Probability. Second edition. Springer, New York 2002.

[4] H. Luckock: A steady-state model of the continuous double auction. Quantitative Finance *3* (2003), 385–404.

[5] S. Maslov: Simple model of a limit order driven market. Physica A *278* (2000), 571–578.

[6] S. Mike and J. D. Farmer: An empirical behavioral model of liquidity and volatility. J. Econom. Dynamics Control *32* (2008), 200–234.

[7] D. Pollard: A User's Guide to Measure Theoretic Probability. Cambridge Univ. Press, Cambridge 2002.

[8] M. Šmíd: On Approximation of Stochastic Programming Problems. PhD Thesis, Charles University, Department of Probability Statistics, 2004.

[9] M. Šmíd: Price tails in the Smith and Farmer's model. Bull. Czech Econometric Soc. *25* (2008), 31–40.

[10] M. Šmíd: Probabilistic Properties of the Continuous Double Auction. Research Report No. 2304. Institute of Information Theory and Automation, Prague 2011.

[11] E. Smith, J. D. Farmer, L. Gillemot, and S. Krishnamurthy: Statistical theory of the continuous double auction. Quantitative Finance *3* (2003), 6, 481–514.

Martin Šmíd, Institute of Information Theory and Automation – Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.

e-mail: smid@utia.cas.cz