

PROBABILISTIC PROPERTIES OF THE CONTINUOUS DOUBLE AUCTION - UNIFORM CASE

Martin Šmíd

Institute of Information Theory and Automation,
Academy of Sciences of the Czech Republic *†
<mailto:smid@utia.cas.cz>

August 5, 2009

Abstract

We study probabilistic properties of a zero intelligence model of a limit order market, very similar to those of [4] and [8]. We (recursively) describe the distributions of the order books and the best quotes. Based on these theoretical results, a procedure for statistical inference of the model may be designed and the evolution of the process may be simulated more efficiently than by the crude simulation of all the events.

Keywords. Continuous double auction, limit order markets, distribution, simulation, statistical inference, price increment tails.

1 Introduction

Recently, several zero-intelligence models of limit order markets¹ have been introduced: [4] shows that even a simple model assuming a Poisson orders' arrival and the uniform distribution of the limit prices generates fat-tailed price increments. [6] computes the rate of the tail by means of a mean-field approximation. [8] introduce a model, similar to that of [4], including, in addition, order cancelations.² A generalization of this model incorporating the statistical properties of real-life order books is made by [5]. Even though many stylized facts cannot be explained by the zero-intelligence models (see [1]) they may be regarded as a good first approach.

Despite the great effort of the authors, no exact probabilistic description of any of the zero-intelligence has been published yet, which, among others, disallows statistical inference of the models. In our recent work [7] we formulated a general model, covering fully the models of [4] and [8] (the latter after a discretization) and partially the one by [5] and we recursively described its distribution. The purpose of the present paper is to apply these results to a simpler (uniform) model of the continuous double auction.

In particular, we consider a model with Poisson order flows, constant cancelation rate and continuous uniform distribution of limit prices (Section 2) which may be easily transformed both to the Maslov's model [4] (by sending the cancelation rate to infinity) and to the model of [8] (by a slight redefinition and the rounding of the prices, see [7], Section 3). The distribution of the present model is described in Section 3). Further (Section 4), we discuss applications of our results. Finally, we conclude the paper (Section 5).

*This work is supported by the grants No. 402/09/0965 and No. 402/07/1113 of the Czech Science Foundation and by project No. LC06075 of the Ministry of Education, Youth and Sports.

†This work is supported by the grants no. 402/06/1417 and 402/07/1113 of the Czech Science Foundation.

¹For a detailed description of a limit order market, see e.g. [8].

²In a subsequent work [2] it is argued that, despite the radical assumption of economic agents acting as gas particles, zero intelligence models replicate several stylized facts found in real-life limit markets.

2 Definitions

2.1 Inputs

In the model we study, the orders of all the four types (i.e. buy/sell market/limit orders) arrive with (possibly different) constant intensities, the intensity of the limit orders' cancelations is constant and the limit prices are uniformly distributed. In particular,

- the arrivals of sell limit orders form a marked Poisson process x with an intensity ι and with marks $\pi_i \sim U(-h, h)$, $u_i \sim \text{Exp}(v)$, standing for the absolute limit price, lifetime of the order respectively, where h is a finite positive constant³ and where all the marks are mutually independent and independent of the arrival times,
- the arrivals of sell market orders form a Poisson process \bar{x} with an intensity $\bar{\iota}$,
- the process y of arrivals of buy limit orders is Poisson with an intensity ω and with independent, mutually independent marks $(r_i, \zeta_i)_{i=1}^{\infty}$ (absolute limit prices, cancelation times respectively) such that $r_i \sim U(h, h)$ and $v_i \sim \text{Exp}(z)$ for some constant z ,
- the process \bar{y} of arrivals of buy market orders is Poisson with an intensity $\bar{\omega}$,
- the random elements x, y, \bar{y}, \bar{x} are mutually independent.

2.2 Dynamics of the System

We describe the state of the market at a time τ by a tuple

$$\Xi_\tau = (A_\tau, B_\tau) \quad (1)$$

where A_τ and B_τ are simple atomic measures (collections of points on the real line) describing the sell order book, buy order book respectively, each atom (point) standing for a waiting limit order with the (absolute) limit price equal to its location. We denote

$$a_\tau \triangleq \min\{\pi : \pi \text{ is an atom of } A_\tau\} \wedge h \quad (2)$$

the value of the (best) ask (we put $\min \emptyset = \infty$) and

$$b_\tau \triangleq \max\{\pi : \pi \text{ is an atom of } B_\tau\} \vee -h \quad (3)$$

the value of the (best) bid (we put $\max \emptyset = -\infty$).

We assume our process to start by a single limit order on each side, i.e. both A_0 and B_0 contain a single (deterministic) point π_0, ρ_0 respectively, such that $\rho_0 < \pi_0$. Further, we assume the lifetime u_0 (v_0) of the starting sell (buy) order to be exponentially distributed with parameter $v, (z)$ such that u_0, v_0 and (x, \bar{x}, y, \bar{y}) are mutually independent.

We let the process Ξ (possibly) jump only at the times $(\tau_i)_{i \in \mathbb{N}}, (\bar{\tau}_i)_{i \in \mathbb{N}}, (\sigma_i)_{i \in \mathbb{N}}, (\bar{\sigma}_i)_{i \in \mathbb{N}}$, denoting the jump times of x, \bar{x}, y , and \bar{y} respectively, or at $(\eta_i)_{i \in \mathbb{N}}, (\zeta_i)_{i \in \mathbb{N}}$ where $\eta_i \triangleq \tau_i + u_i$ and $\zeta_i \triangleq \sigma_i + v_i$, $i \geq 0$ (i.e. the cancelation times of sell limit orders, buy limit orders respectively, we put $\tau_0 = \sigma_0 = 0$).

If τ is one of the possible jump times of Ξ then the change of Ξ at τ is defined as follows:

- if $\tau = \tau_i$ for some $i > 0$ and $\pi_i > b_{\tau-}$ then π_i is added into A
- if $\tau = \eta_i$ for some $i \geq 0$ and π_i is present in $A_{\tau-}$ then π_i is removed from A
- if $\tau = \bar{\tau}_i$ for some $i > 0$ and $B_{\tau-} \neq 0$ then $b_{\tau-}$ is removed from B
- if $\tau = \sigma_i$ for some $i > 0$ and $\rho_i < a_{\tau-}$ then ρ_i is added into B
- if $\tau = \zeta_i$ for some $i \geq 0$ and ρ_i is present in $B_{\tau-}$ then ρ_i is removed from B
- if $\tau = \bar{\sigma}_i$ for some $i > 0$ and $A_{\tau-} \neq 0$ then $a_{\tau-}$ is removed from A
- if τ does not fit any of the conditions above then both A and B are left unchanged at τ .

³Later we define our model also for infinite h

It follows from the definition of the input processes that

$$\mathbb{P}[\text{any pair of } \tau_i, \bar{\sigma}_i, \sigma_i, \bar{\tau}_i, \eta_i, \zeta_i, i \in \mathbb{N}, \text{ coincide}] = 0 \quad (4)$$

hence our definition is correct with probability one. Further, since

$$\mathbb{P}[\text{any pair of } \pi_i, \rho_i, i \in \mathbb{N}, \text{ coincide}] = 0 \quad (5)$$

it is guaranteed that A_τ and B_τ are simple (with no overlapping points) at each $\tau \geq 0$.

For a “formula based” definition of Ξ , see [7].

3 Distributions

In the present Section, we describe probabilistic properties of Ξ and of the process of the best quotes

$$\xi_\tau \triangleq (a_\tau, b_\tau). \quad (6)$$

An uninterested reader may skip this section; however, (s)he will need to return here for a notation sometimes.

We start with a nearly obvious statement:

Proposition 1 (*Markov properties of Ξ*)

(i) For any deterministic $0 \leq s_1 < \dots < s_n < s$,

$$\mathcal{L}(\Xi_{s+\bullet} | \Xi_s, \Xi_{s_1}, \dots, \Xi_{s_n}) = \mathcal{L}(\Xi_{s+\bullet} | \Xi_s) \quad (7)$$

(here, for any random elements X, Y , symbol $\mathcal{L}(X|Y)$ denotes the conditional distribution of X given Y).

(ii) Relation (7) keeps holding even if $s_1 < \dots < s_n < s$ are optional times with respect to the filtration generated by Ξ .

Proof. Statement (i) follows from the fact that all the inter-jump times and the orders’ lifetimes are exponential and that the value of Ξ at a jump is fully determined by the value at the last jump and the type of jump time. Part (ii) stems from the fact that Ξ is pure jump type process (see [3], chp. 12). \square

To go on, we need to introduce some notation: For any interval \mathfrak{J} such that $\mathfrak{J} = [s, t)$ or $\mathfrak{J} = [s, t]$, $0 \leq s \leq t$, denote

$$\tilde{a}^\mathfrak{J} : \mathfrak{J} \rightarrow \mathbb{R}, \quad \tilde{a}_\tau^\mathfrak{J} \triangleq \max_{\theta \in \mathfrak{J} \cap [\tau, \infty)} a_\theta, \quad \tau \in \mathfrak{J}, \quad (8)$$

and introduce a function

$$\kappa^\mathfrak{J}(\mathbf{p}) \triangleq \sum_{j=0}^{J^\mathfrak{J}} \frac{\iota^{\star}}{v} [\mathbf{p} \wedge \tilde{a}_{[j-1]}^\mathfrak{J} - \tilde{a}_{[j]}^\mathfrak{J}]^+ [1 - \exp\{-v(t - \varsigma_j^\mathfrak{J})\}], \quad \mathbf{p} \in \mathbb{R}, \quad (9)$$

where $\iota^{\star} \triangleq (2h)^{-1}\iota$, $J^\mathfrak{J}$ is the number of jumps of $\tilde{a}^\mathfrak{J}$ on \mathfrak{J} , $\varsigma_1^\mathfrak{J} < \dots < \varsigma_{J^\mathfrak{J}}^\mathfrak{J}$ denote the jump times themselves (we put $\varsigma_0^\mathfrak{J} \triangleq s$) and where

$$\tilde{a}_{[j]}^\mathfrak{J} \triangleq \tilde{a}_{\varsigma_j^\mathfrak{J}}^\mathfrak{J}, \quad 0 \leq j \leq J^\mathfrak{J}, \quad (10)$$

(we put $\tilde{a}_{[-1]}^\mathfrak{J} \triangleq h$). Further, denote $0 = \vartheta_0 < \vartheta_1 < \dots$ the sequence of jump times of ξ .

3.1 Order Books

Until the end of subsection 3.1, fix $s \leq t$ fulfilling one of the conditions

- both s and t are deterministic,
- $s = \vartheta_k$ and $t = \vartheta_i$ for some (deterministic) $k, i \in \mathbb{N}$, $k \leq i$,

and agree to write \tilde{a} instead of $\tilde{a}^{[s,t]}$, J instead of $J^{[s,t]}$, etc.

The following two Propositions describe the distribution of the order books given $\Xi_s, \xi_{[s,t]}$ where the latter symbol denotes the trajectory of ξ restricted to interval $[s, t]$.

Proposition 2 (conditional distribution of A_t)

$$A_t = \delta_{a_t} + \sum_{j=1}^J \delta_{\tilde{a}_{[j]}} e_j + \sum_{j=1}^K \delta_{\alpha_j^s} d_j + L. \quad (11)$$

where δ_q denotes the Dirac measure concentrated in q (i.e. a single point $\{q\}$) and

- e_1, \dots, e_J are binary variables such that

$$\mathcal{L}(e_j | \Xi_s, \xi_{[s,t]}) = \text{Alternative}(\exp\{-v(t - \varsigma_j)\}), \quad 1 \leq j \leq J, \quad (12)$$

- $\alpha_1 < \dots < \alpha_K$ are all the atoms of A_s whose location is greater than \tilde{a}_s ,

- d_1, \dots, d_K are binary variables with

$$\mathcal{L}(d_j | \Xi_s, \xi_{[s,t]}) = \text{Alternative}(\exp\{-v(t - s)\}), \quad 1 \leq j \leq K, \quad (13)$$

- L is (conditionally on $\Xi_s, \xi_{[s,t]}$) a non-homogenous Poisson process on (a_t, h) with the intensity given by distribution function κ ,

- $e_1, \dots, e_J, d_1, \dots, d_K, L$ are conditionally independent given $(\Xi_s, \xi_{[s,t]})$.

Proof. See [7]. \square

Remark 1 Using the symmetry of A and B , an analogous formula for the distribution of B may be obtained as a corollary of Proposition 2.

3.2 Best Quotes

Until the end of subsection 3.2, fix deterministic integers $0 \leq k < i$ and agree to abbreviate $\tilde{a}^{[\vartheta_k, \vartheta_i]}$ as \tilde{a} , $J^{[\vartheta_k, \vartheta_i]}$ as J etc.

Our present goal is to specify the conditional distribution of ϑ_{i+1} and $\xi_{\vartheta_{i+1}}$ given Ξ_{ϑ_k} and $\xi_{[\vartheta_k, \vartheta_i]}$. For a better intuitive understanding, we add an additional step to the definition: we consider an supplementary variable χ_{i+1} , coding the type of the event happening at the time ϑ_{i+1} . Denoting the possible values of χ_{i+1} symbolically by \mathbf{a}^+ , \mathbf{a}^- , \mathbf{b}^+ and \mathbf{b}^- , we define it as

$$\chi_i \triangleq \begin{cases} \mathbf{a}^+ & \text{if } \Delta a_{\vartheta_i} > 0, \Delta b_{\vartheta_i} = 0, \\ \mathbf{a}^- & \text{if } \Delta a_{\vartheta_i} < 0, \Delta b_{\vartheta_i} = 0, \\ \mathbf{b}^+ & \text{if } \Delta a_{\vartheta_i} = 0, \Delta b_{\vartheta_i} > 0, \\ \mathbf{b}^- & \text{if } \Delta a_{\vartheta_i} = 0, \Delta b_{\vartheta_i} < 0. \end{cases} \quad (14)$$

Denote

$$\varsigma_i^{\mathbf{a}^-} = \iota^*(a_{\vartheta_i} - b_{\vartheta_i}), \quad \varsigma_i^{\mathbf{a}^+} = \mathbf{1}_{[a_{\vartheta_i} < h]}(\bar{\omega} + v), \quad (15)$$

$$\varsigma_i^{\mathbf{b}^-} = \mathbf{1}_{[b_{\vartheta_i} > -h]}(\bar{\iota} + z), \quad \varsigma_i^{\mathbf{b}^+} = \omega^*(a_{\vartheta_i} - b_{\vartheta_i}), \quad (16)$$

$$\varsigma_i = \varsigma_i^{\mathbf{a}^+} + \varsigma_i^{\mathbf{a}^-} + \varsigma_i^{\mathbf{b}^+} + \varsigma_i^{\mathbf{b}^-}, \quad \omega^* \triangleq (2h)^{-1}\omega, \quad (17)$$

(ι_x^* was introduced at the definition of $\kappa^{\mathcal{J}}$).

Proposition 3 (distribution of ϑ_{i+1})

$$\mathcal{L}(\Delta\vartheta_{i+1} | \xi_{[\vartheta_k, \vartheta_i]}, \Xi_{\vartheta_k}) = \text{Exp}(\varsigma_i). \quad (18)$$

Proposition 4 (distribution of χ_{i+1})

$$\mathcal{L}(\chi_{i+1} \mid \Delta\vartheta_{i+1}, \xi_{[\vartheta_k, \vartheta_i]}, \Xi_{\vartheta_k}) = \left(\frac{\varsigma_i^{a^+}}{\varsigma_i}, \frac{\varsigma_i^{a^-}}{\varsigma_i}, \frac{\varsigma_i^{b^+}}{\varsigma_i}, \frac{\varsigma_i^{b^-}}{\varsigma_i} \right). \quad (19)$$

Proposition 5 (distribution of $a_{\vartheta_{i+1}}$)

(i) If $\chi_{i+1} = a^+$ then

$$\begin{aligned} \mathbb{P}[a_{\vartheta_{i+1}} > \mathfrak{p} \mid \Delta\vartheta_{i+1}, \chi_{i+1}, \xi_{[\vartheta_k, \vartheta_i]}, \Xi_{\vartheta_k}] \\ = \mathbf{1}_{[\mathfrak{p} < h]} \exp\{-\kappa(\mathfrak{p})\} (1 - \exp\{-v(\vartheta_{i+1} - \vartheta_k)\})^{K(\mathfrak{p})} \\ \times \prod_{1 \leq j < J, \tilde{a}_{[j]} \leq \mathfrak{p}} (1 - \exp\{-v(\vartheta_{i+1} - \varsigma_{j+1})\}) \end{aligned}$$

where $K(\mathfrak{p})$ is the number of points of A_{ϑ_k} belonging to interval $(\tilde{a}_{\vartheta_k}, \mathfrak{p}]$,

(ii) If $\chi_{i+1} = a^-$ then

$$\mathcal{L}(a_{\vartheta_{i+1}} \mid \Delta\vartheta_{i+1}, \chi_{i+1}, \xi_{[\vartheta_k, \vartheta_i]}, \Xi_{\vartheta_k}) = \mathbf{U}(b_{\vartheta_i}, a_{\vartheta_i}) \quad (20)$$

(iii) If $\chi_{i+1} \in \{b^-, b^+\}$ then

$$\mathcal{L}(a_{\vartheta_{i+1}} \mid \Delta\vartheta_{i+1}, \chi_{i+1}, \xi_{[\vartheta_k, \vartheta_i]}, \Xi_{\vartheta_k}) = \delta_{a_{\vartheta_i}}. \quad (21)$$

Proof of Propositions 3-5. See [7]. \square

Remark 2 The formula for the distribution of $b_{\vartheta_{i+1}}$ is symmetric.

4 Applications

Even if the formulas describing the distribution are quite complicated, several useful and practical applications may be constructed based on them.

4.1 Statistical inference of the model

The problem of the inference of our model (i.e. testing its validity and estimating its parameters) may be broken into five parts: an inference of all the four input processes. Since the flows of market orders may be studied by standard techniques,⁴ we do not discuss them in the present work. To test and infer the distributions of $(\tau_\nu)_{\nu \in \mathbb{N}}$ and $(p_\nu)_{\nu \in \mathbb{N}}$, may also apply standard statistical tools. On the other hand, we are getting into difficulties with cancellations rates v and z because the sample of observed cancellations of the orders is censored (some orders are executed before they could be canceled) - in fact, this is the point where our theoretical results may help: based on our formulas, the cancellation rates (together with the market limit orders' arrival rates, if we are lacking detailed enough data to infer them directly) may be estimated. For details, see Section 4.2. of [7].

4.2 Efficient Simulation

Since we do not know analytic formulae for the (unconditional) distributions of Ξ and ξ , a Monte Carlo simulation is needed when working with the processes. However, our knowledge of the (conditional) distributions may help us to speed up the simulations significantly.

Contrary to the “crude” simulation, when each variable defining the system is generated, our procedure allows us to omit a vast majority of the steps of such a simulation: Thanks to our results, it suffices to generate the process ξ only and to draw the values of Ξ only at the times when they are actually needed.⁵ Our procedure is as follows:

⁴Of course, this is true only if we are able to distinguish the moves of bid, ask respectively, caused by cancellations from those due to market orders.

⁵In fact, our procedure “is able” to generate the order books only at the times of jumps of ξ ; however, if we wanted to draw all the values of the order books, we may use the crude simulation since the last jump of ξ .

For each $i \in \mathbb{N}_0$:

1. Generate $\Delta\vartheta_{i+1}$ and χ_{i+1} from their conditional distribution given the $(\xi_{[s_i, \vartheta_i]}, \Xi_{s_i})$ where s_i is the time of the last generation of a value of Ξ (by Proposition 3 and 4)
2. Generate $a_{\vartheta_{i+1}}$ and $b_{\vartheta_{i+1}}$ from their conditional distributions given $(\chi_{i+1}, \Delta\vartheta_{i+1}, \xi_{[s_i, \vartheta_i]}, \Xi_{s_i})$ (by Proposition 5 and its symmetric counterpart)
3. If needed, generate $B_{\vartheta_{i+1}}$ and/or $A_{\vartheta_{i+1}}$ from their conditional distributions given $(\xi_{[s_1, \vartheta_{i+1}]}, \Xi_{s_i})$ (by Proposition 2 and its symmetric counterpart).

5 Conclusion

We examined theoretically some zero-intelligence models of limit order markets and showed some applications of our results. In particular, we outlined a way of a rigorous statistical inference of such models and we proposed a way of an efficient Monte Carlo simulation of the models. Further generalizations of the model are subjects of our future research.

References

- [1] J. P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of ‘random’ price changes. *Quantitative Finance*, 4(2):176–190, 2004. Preprint at <http://www.science-finance.fr/papers/0307332.pdf>.
- [2] J. D. Farmer, P. Patelli, and I. I. Zovko. The predictive power of zero intelligence in financial markets. *Proc. Nat. Acad. Sci.*, 102:2254–2259, 2005. Preprint at <http://www.pnas.org/cgi/content/full/102/6/2254>.
- [3] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, second edition, 2002.
- [4] S. Maslov. Simple model of a limit order driven market. *Physica A*, 278:571–578, 2000.
- [5] S. Mike and J. D. Farmer. An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics & Control*, 32(1):200–234, 2008.
- [6] F. Slanina. Mean-field approximation for a limit order driven market. *Phys. Rev.*, 64:056136, 2001. Preprint at <http://xxx.sissa.it/abs/cond-mat/0104547>.
- [7] M. Šmíd. Probabilistic properties of the continuous double auction, 2008. Submitted to *Kybernetika*, Available at SSRN: <http://ssrn.com/abstract=1107114>.
- [8] E. Smith, J. D. Farmer, L. Gillemot, and S. Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative finance*, 3(6):481–514, 2003. <http://www.santafe.edu/~baes/jdf/papers/doubleauctionfinal.pdf>.