

# LOW-TRAFFIC LIMIT AND FIRST-PASSAGE TIMES FOR A SIMPLE MODEL OF THE CONTINUOUS DOUBLE AUCTION

ENRICO SCALAS, FABIO RAPALLO, AND TIJANA RADIVOJEVIĆ

ABSTRACT. We consider a simplified model of the continuous double auction where prices are integers varying from 1 to  $N$  with limit orders and market orders, but quantity per order limited to a single share. For this model, the order process is equivalent to two  $M/M/1$  queues. We study the behaviour of the auction in the low-traffic limit where limit orders are immediately transformed into market orders. In this limit, the distribution of prices can be computed exactly and gives a reasonable approximation of the price distribution when the ratio between the rate of order arrivals and the rate of order executions is below  $1/2$ . This is further confirmed by the analysis of the first passage time in 1 or  $N$ .

## 1. INTRODUCTION

Most of the regulated markets in the world implement a trading mechanism known as the *continuous double auction* to match supply and demand. This mechanism has two sides. On the supply side there are orders to sell and on the demand side there are orders to buy. Hence, the auction is called *double*. Moreover it occurs in continuous time. Hence, it is called *continuous*.

In recent years, the theory of this auction has gained more and more interest. In particular, it has been shown that appropriate models of the double auction can be mapped in a multi-class queue [1], so that its ergodic properties and the limiting invariant distribution can be studied using established techniques [2].

In this paper, we consider a simplified model (see [3] and references therein) where prices take  $N$  integer values from 1 to  $N$ . Only two types of orders are considered: *limit orders* and *market orders*. In their turn, limit orders can be either orders to sell a single share at a price not lower than a given amount (*asks*) or orders to buy a single share at a price not higher than a given amount (*bids*). In other words, the quantity attached to every limit order is always 1. Among all the asks, the *best ask* is the smallest ask price, whereas the *best bid* is the largest bid price. The best bid is always strictly smaller than the best ask. Market orders have also two sides: either they accept the available best bid or the available best ask. For the sake of simplicity, limit ask orders and limit bid orders arrive according to a Poisson process at a rate  $\lambda_a$  and  $\lambda_b$ , respectively. In the following,

we assume symmetry, i.e.  $\lambda_a = \lambda_b = \lambda$ . Market orders to buy and market orders to sell arrive separated by durations following the exponential distribution with parameter  $\mu_b$  and  $\mu_a$ , respectively. Again, symmetry is assumed, namely  $\mu_a = \mu_b = \mu$ . Limit ask orders follow the uniform distribution in the interval from  $p_b + 1$  to  $p_b + n$ , where  $p_b$  is the current best bid and  $n \geq 1$  is a parameter of the model. Similarly, limit bid orders are uniformly drawn from the interval  $p_a - n$  to  $p_a - 1$ , where  $p_a$  is the the current best ask. The accessible states of the auction are limited by the condition  $p_b < p_a$ . When  $p_a$  is between 1 and  $n$  ( $p_b$  between  $N - n + 1$  and  $N$ ), the bid (respectively ask) interval is restricted correspondingly. For instance, if  $p_a = 1$ , bids are impossible. The parameter  $n$  acts as a cut-off for price jumps. Eventually, if no orders are present in the auction, the next bid,  $b$ , is uniformly chosen in the interval  $p - n \leq b \leq p$  and the next ask,  $a$ , is uniformly taken from  $p \leq a \leq p + n$ , where  $p$  is the price of the last trade. Specifying an initial price (the opening auction price) is sufficient to start the auction. A short remark is necessary at this stage: It turns out that order inter-arrival times are not exponentially distributed in real markets (see [4] and references therein). This means, that the above description in terms of  $M/M/1$  processes should be replaced by a semi-Markov description in terms of  $G/G/1$  processes. However, in this paper, for the sake of simplicity, we will limit our analysis to the Markovian case.

The model described above is essentially the same as in [5] and in [6]. It is a *zero intelligence agent-based model* [7]. As already mentioned in [3], this version of the model does not use the uniform distribution over  $[0, \infty)$  as in [5] and it is not limited to the case in which limit orders arrive only at the best bid/ask price as in [6]. A preliminary discussion of this version was presented in [8]. This model was extensively studied in [9], in the case in which price movements equal one tick. These authors also studied the heavy-traffic limit [10] where functional limit theorems are available leading to diffusion approximations [11, 12].

In [3], the focus was on the ergodic properties of the model. Based on the fact that the order process is equivalent to two independent  $M/M/1$  queues, it was shown that there are three regimes depending on the value of  $\rho = \lambda/\mu$ . For  $0 < \rho < 1$ , prices are free to fluctuate over the full price range and statistical equilibrium is reached (ergodic regime). For  $\rho \geq 1$ , the auction is in a non-ergodic regime which stabilizes prices. Due to the presence of the parameter  $n$ , there is an additional transition. If  $1 \leq \rho < n$ , prices can still fluctuate in a limited range, whereas for  $\rho \geq n$ , prices eventually fluctuate between two values. This regime cannot be found if one only considers the case  $n = 1$ .

In the following, we further characterize the ergodic regime by considering the so-called *low-traffic* limit where  $\rho \ll 1$ . It is a limit where analytic results are

available for the price dynamics as discussed below. Moreover, we study the first-passage time of the auction in 1 or in  $N$ . It turns out that this analysis provides useful approximations for the behaviour of the auction when  $\rho < 1/2$ .

## 2. THE LOW-TRAFFIC LIMIT

In the low-traffic limit ( $\rho \ll 1$ ), when limit orders arrive, they are immediately transformed into market orders. The book is almost always empty. In this limit, it is possible to explicitly write the transition probabilities for the price process and study the price Markov chain for any value of  $n$  and  $N$ . To give an idea on how to proceed, let us assume that the initial price is  $p$ . Then, the conditional probability of a bid is given by

$$(1) \quad \mathbb{P}(B_1 = b | P_0 = p) = \begin{cases} 0 & \text{if } b < p - n \text{ or } b > p \\ \frac{1}{p} & \text{if } 1 \leq b \leq p \leq n \\ \frac{1}{n+1} & \text{if } 1 \leq p - n \leq b \leq p. \end{cases}$$

This bid is immediately accepted and it becomes the next price. A similar set of equations can be written for the asks conditioned to the initial price.

$$(2) \quad \mathbb{P}(A_1 = a | P_0 = p) = \begin{cases} 0 & \text{if } a < p \text{ or } a > p + n \\ \frac{1}{N-p+1} & \text{if } N - n + 1 \leq p \leq a \leq N \\ \frac{1}{n+1} & \text{if } p \leq a \leq p + n \leq N. \end{cases}$$

Both equations are an immediate consequence of the model definition. For a full characterization of the price Markov chain, the distribution of the initial price is needed. For instance, if the initial price is chosen uniformly, the probability of an initial price is  $1/N$ ; if the chain starts from a given price, the probability of this price is 1 and the probabilities of all the other prices are 0, and so on. In a symmetric auction, for which  $\lambda_a = \lambda_b = \lambda$  and  $\mu_a = \mu_b = \mu$ , the probability of a bid arriving is  $1/2$  and it is equal to the probability of arrival of an ask. Therefore, in the low traffic regime, the transition probability for prices is given by

$$(3) \quad P_{p,p'} = \mathbb{P}(P_1 = p' | P_0 = p) = \frac{1}{2} \mathbb{P}(A_1 = p' | P_0 = p) + \frac{1}{2} \mathbb{P}(B_1 = p' | P_0 = p).$$

Let us now assume that the limit  $\rho \ll 1$  is realized by keeping the arrival rate  $\lambda$  finite and letting  $\mu \gg \lambda$ . Then, if  $N(t)$  denotes the number of transactions up to time  $t$ , we have that  $N(t)$  is Poisson distributed with parameter  $2\lambda$ , given that the auction has two sides (either a bid arrives with rate  $\lambda$  or an ask arrives with rate  $\lambda$ ). In fact,  $N(t)$  is the superposition of two Poisson processes with parameter

$\lambda$ . In other words, the price process can be seen as an embedded Markov chain characterized by the transition probability (3) subordinated to the Poisson process  $N(t)$ . Once this remark is made, it is safe to focus on the embedded chain and study its properties. In particular we are interested in the convergence of the price probability. First of all, we notice that after any transactions, the double auction is exactly in the same situation as in the initial case, except for the fact that the price probability varies with time. In other words, the Markov chain defined above is homogeneous. From the study of the transition probability, one can further infer that the Markov chain is irreducible. In fact, it is possible to reach any price from any other price. Moreover, given that the diagonal terms of the Markov transition matrix are all positive, meaning that there is a finite probability for the price not to change at every step, we can conclude that our Markov chain is aperiodic. Being irreducible and aperiodic, our chain has a unique invariant distribution and this is an equilibrium distribution.

In order to illustrate the above findings, let us consider a specific example with  $N = 10$  prices and  $n = 2$ . In this case, the price transition probability matrix is

$$(4) \quad P = \begin{pmatrix} 4/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 5/12 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 5/12 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 4/6 \end{pmatrix}.$$

The invariant distribution is obtained by looking for the left eigenvector with unit eigenvalue, namely

$$(5) \quad \pi P = \pi,$$

which, in the case of (4) gives

$$(6) \quad \pi = (0.1171, 0.0895, 0.1, 0.0961, 0.0974, 0.0974, 0.0961, 0.1, 0.0895, 0.1171).$$

In the Appendix A, we present a general algorithm to find the invariant distribution of prices. For  $n = 1$ , there is a remarkable result. In fact, in this case, the transition matrix is a symmetric, doubly-stochastic matrix. Since  $P\mathbf{1} = \mathbf{1}$  and  $\mathbf{1}^T P = \mathbf{1}^T$  (because row sums and column sums are 1), then

$$(7) \quad \frac{1}{N}\mathbf{1}^T P = \frac{1}{N}\mathbf{1}^T,$$

and the uniform distribution is the invariant distribution for the Markov chain.

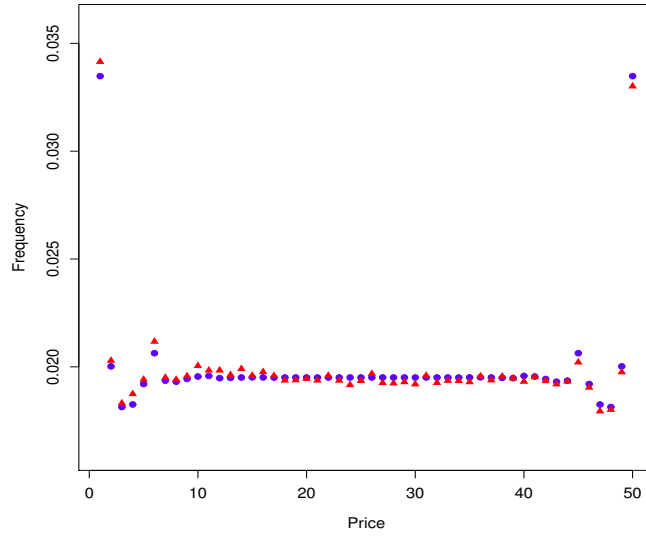


FIGURE 1. Equilibrium price distribution in the low-traffic limit in the case  $N = 50$ ,  $n = 5$  (circles). The triangles denote the price frequency for a Monte Carlo simulation of the double auction with  $\rho = 10^{-4}$  after  $10^6$  steps.

Always for purpose of illustration, in Figure 1, we plot the low-traffic limit price distribution for the case  $N = 50$ ,  $n = 5$  and we compare it with the frequency with which states appear after equilibration in a Monte Carlo simulation of the chain after  $10^6$  iterations for  $\rho = 10^{-4}$ . In this case  $10^6$  iterations are already sufficient to show that the agreement between the low-traffic-limit approximation and the result of Monte Carlo simulations is good. It is striking to observe that the price distribution in the low-traffic limit is still a reasonable approximation when  $\rho = 0.3$  as shown in Figure 2. The approximation breaks down for  $\rho \geq 0.5$  as shown in Figure 3 for the case  $\rho = 0.9$ . This behaviour of the price distributions leads to a different behaviour for first passage times at the boundary prices. In fact, for  $\rho > 0.5$  and  $n > 1$  the residence time of the systems close to the boundaries becomes negligible as shown by Figure 3, leading to an increase of the value of the average first passage time.

### 3. FIRST PASSAGE TIMES

In this section we shall focus on first passage times. Given a double auction with  $N$  possible prices, labeled with the integers  $1, \dots, N$ , we fix the initial price at the median point  $\lfloor \frac{N+1}{2} \rfloor$ , where  $\lfloor \cdot \rfloor$  denotes the floor operator. We study the random variable  $T$ : The first passage time at 1 or at  $N$ ; for this reason, our problem belongs

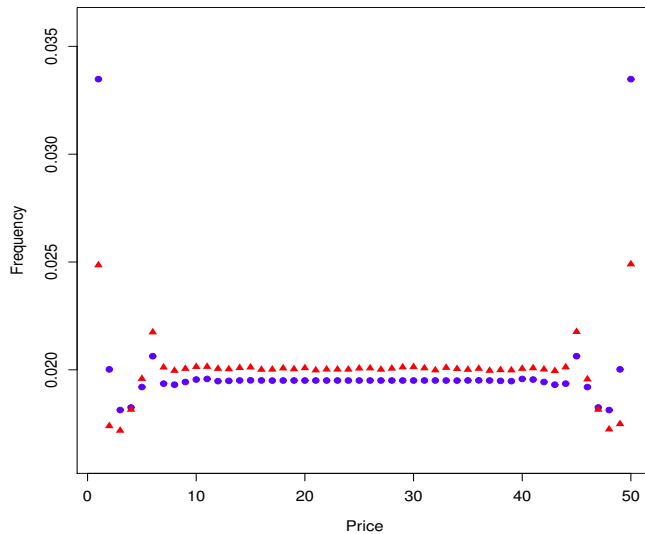


FIGURE 2. Comparison between the equilibrium price distribution in the low-traffic limit in the case  $N = 50$ ,  $n = 5$  (circles) and a Monte Carlo simulation (triangles) of the double auction with  $\rho = 0.3$  after  $10^9$  steps.

to the class of two-barrier problems, and, for simplicity, we will assume that the price walk is symmetric. The behaviour of the first passage time distribution has been studied in several similar problems, both with theoretical results on the exact or asymptotic distribution, and through simulation studies, see e.g. [14] and [15].

Here, we present the results of a simulation study to investigate the main features of the distribution of  $\log(T)$ . In particular, we compare such a distribution with the theoretical distribution derived under the low traffic assumption. Since the distribution of  $T$  is highly skewed with a fat upper tail, as shown in figures, all the plots reported here refer to the distribution of its natural logarithm  $\log(T)$ . For sake of simplicity, we have performed all comparisons in this section with the parameter  $\mu = 1$  fixed. With this assumption, we get the low traffic limit if  $\rho = \lambda \ll 1$ .

In Figure 4 the histograms of  $\log(T)$  for  $n = 5$  and for  $\rho = 0.02$  are displayed for 4 different values of  $N$ , namely  $N = 10, 40, 70, 100$ . One can observe that the shape of the distribution is skewed for small values of  $N$ , while it approaches a Gaussian distribution in the case  $N = 100$  (the best-fit normal curve is plotted together with the histogram). The simulations have been implemented in R [13], and all histograms in this study are based on 10,000 Monte-Carlo replicates.

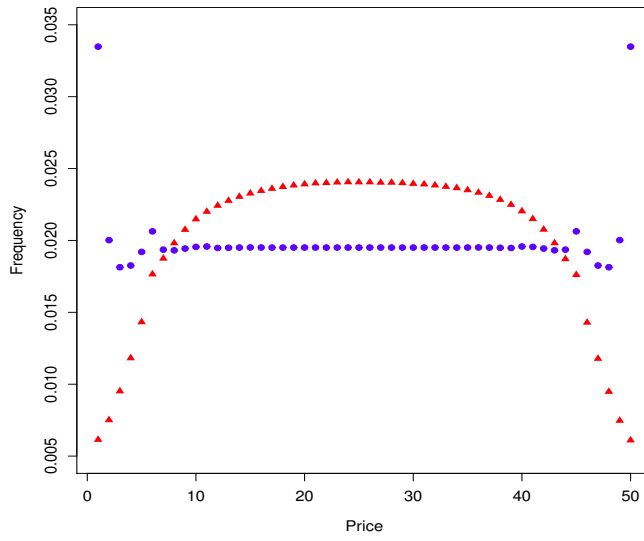


FIGURE 3. Comparison between the equilibrium price distribution in the low-traffic limit in the case  $N = 50$ ,  $n = 5$  (circles) and a Monte Carlo simulation (triangles) of the double auction with  $\rho = 0.9$  after  $10^9$  steps.

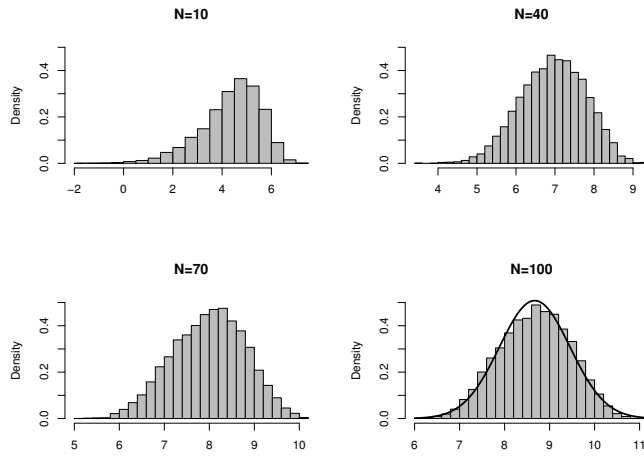


FIGURE 4. Distribution of  $\log(T)$  for  $n = 5$  fixed and  $\rho = 0.02$ .

Figure 5 refers to the same settings as above, but with the ratio fixed at  $\rho = 0.5$ . The distribution of  $\log(T)$  has almost the same shape for  $\rho = 0.5$  and for  $\rho = 0.02$ . In agreement with the conclusions of the previous section, this suggests that the

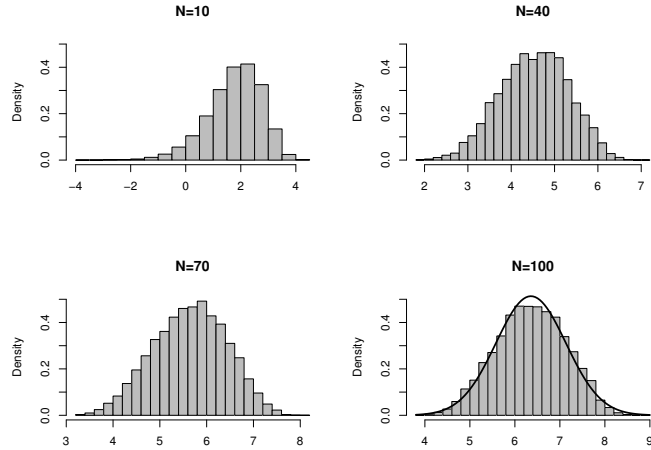


FIGURE 5. Distribution of  $\log(T)$  for  $n = 5$  fixed and  $\rho = 0.5$ .

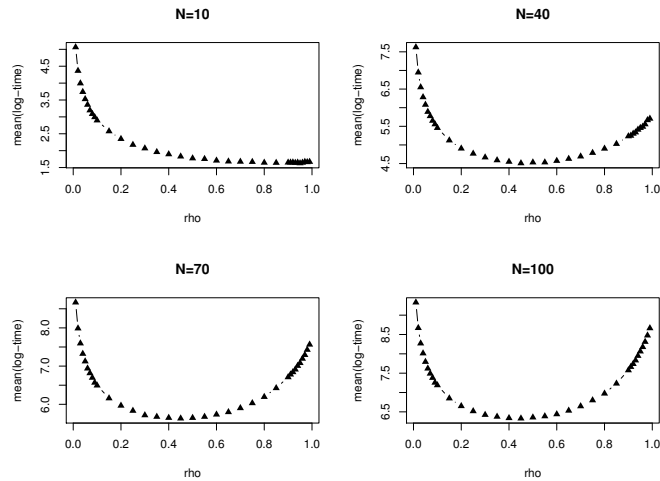


FIGURE 6. Mean of  $\log(T)$  for  $n = 5$  and  $N = 10, 40, 70, 100$ .

behaviour of the low traffic limit is a good approximation also for values of  $\rho$  up to 0.5 also in terms of the first passage time distribution.

Finally, in Figure 6, the means of  $\log(T)$  as a function of  $\rho$  under various choices of  $N$  ( $n = 5$  fixed) are displayed. A minimum occur between  $\rho = 0.4$  and  $\rho = 0.5$  except for the first experimental setting ( $N = 10$ ).

To complete the simulation study, we have compared the distribution of  $T$  with an approximation suggested by the results of the previous section. Basically, we adapt here a known formula for a discrete two-barrier problem. Such formula gives the



distribution of the number of price changes needed to reach the boundary. Then, the parameter  $\rho$  controls the proportion of orders leading to a price change, and therefore it defines the discrete time distribution of the number of orders needed to reach the boundary. Remember that in the low traffic limit when limit orders arrive, they are immediately transformed into market orders. Finally, we add suitable exponential distributions to switch to the continuous framework of our model. To avoid problems in some formulae, we assume here  $N$  to be odd, so that  $(N + 1)/2$  is always integer. We have limited our study to the case  $n = 1$  in order to avoid further technicalities in the formulae and to capture the major features of the model. The approximation is built up in three steps, as detailed below:

- First, consider a discrete first passage time  $\tilde{T}^{(d)}$  in a simple symmetric random walk with two reflecting barriers and discrete  $\pm 1$  steps, whose distribution is

$$(8) \quad \mathbb{P}(\tilde{T}^{(d)} = \tilde{h}) = \frac{2}{N-1} \sum_{k=1}^{N-2} (-1)^{k+1} \sin\left(\frac{k\pi}{N-1}\right) \cos^{\tilde{h}-1}\left(\frac{k\pi}{N-1}\right) \sin\left(\frac{k\pi}{2}\right)$$

for  $\tilde{h} \geq 1$ . The above distribution can be found in [16] and is extensively discussed with several generalizations in [17].

- The rate of arrival of limit orders over all orders is  $\lambda/(\lambda + \mu) = \rho/(\rho + 1)$ , and in the low traffic approximation all limit orders arrive when the book is empty. A limit order to buy (resp. to sell) fixes the price at the old price  $p$  or at  $p + 1$  (resp.  $p - 1$ ) with probability  $1/2$  each. Therefore, at any given time, the price changes with rate  $\frac{\rho}{2(\rho+1)}$ . Thus, given  $\tilde{T}^{(d)} = \tilde{h}$ , consider a Negative Binomial variable  $NB$  with parameters  $\tilde{h}$  and  $\frac{\rho}{2(\rho+1)}$ , and define  $T^{(d)} = NB + 1$ . The variable  $T^{(d)}$  is again a discrete random variable and it counts the number of events after the change of price is actually performed;
- The interarrival time between two consecutive events follows an exponential distribution with mean  $2\mu(1 + \rho)$ . Therefore, given  $T^{(d)} = h$ , the first passage time is approximated by a random variable following a Gamma distribution with parameters  $h$  and  $\frac{1}{2\mu(1+\rho)}$ .

In conclusion, the distribution of the first passage time can be approximated by a suitable mixture  $T^{(a)}$  of Gamma distributions, whose parameters are computed according to the formula in Eq. (8) for the discrete case. Notice that in the previous construction the low-traffic hypothesis is used only when we assume the book empty when a limit order arrives.

To show that this approximation works well for small values of  $\rho$ , we have plotted in Figure 7 the empirical cumulative distribution function (ECDF) of  $\log(T)$  and (a

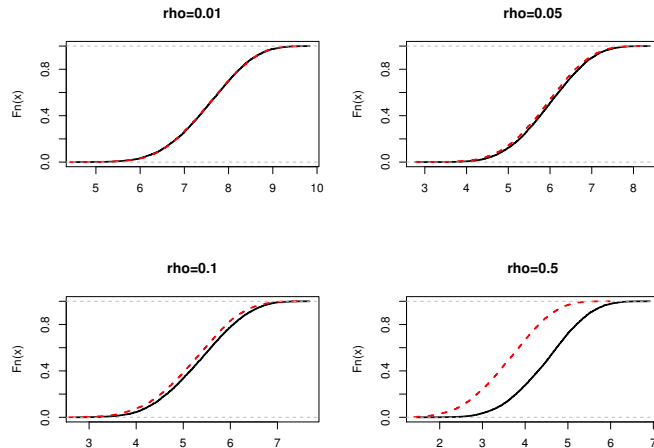


FIGURE 7. ECDFs of  $\log(T)$  (in black) and its low traffic approximation  $\log(T^{(a)})$  (dashed, in red) for  $N = 11$ ,  $n = 1$ .

Monte-Carlo approximation of) the distribution of  $\log(T^{(a)})$  for 4 different values of  $\rho$  in the case  $N = 11$ . Also in these simulations we have considered only the case  $\mu = 1$ .

We can observe in Figure 7 that for  $\rho = 0.01$ ,  $\rho = 0.05$  the simulated distribution and its theoretical approximation are nearly identical (the  $p$ -value of the Kolmogorov-Smirnov based on 10,000 Monte Carlo replicates is 0.7212 for  $\rho = 0.01$  and 0.0116 for  $\rho = 0.05$ ). When  $\rho = 0.1$  the two distributions show some discrepancies, while in the last scenario ( $\rho = 0.5$ ) the approximation fails. The low traffic approximation  $T^{(a)}$  tends to underestimate the distribution of  $T$ . This behaviour is observed also in the  $\rho = 0.1$  case, but it is clearer in the  $\rho = 0.5$  case, as expected.

When  $n > 1$  a formula like that in Eq. (8) is no longer available. However, we can analyze the low-traffic approximation by studying the expected values of the distributions. In fact, the expected value of the first passage time  $\mu_T^{(d)}$  of the discrete chain in the low-traffic approximation can be computed through the linear system

$$(9) \quad (I - P_{2,N-1})x = 1$$

where  $P_{2,N-1}$  is the transition matrix restricted to the transient states,  $I$  is the  $(N-2) \times (N-2)$  identity matrix, and  $1$  is a column vector of 1 with dimension  $N-2$  (see e.g. [18] for details). Then, the mean time in the continuous setting  $\mu_T$  is obtained by scaling  $\mu_T^{(d)}$  by a factor  $1/(2\rho)$ , following the same reasoning as above. In Table 1, the means  $\bar{T}$  of the Monte Carlo simulations and the theoretical expected value

$N$	$n$	$\rho = 0.01$			$\rho = 0.02$			$\rho = 0.05$		
		$\bar{T}$	$\mu_T$	$\Delta\%$	$\bar{T}$	$\mu_T$	$\Delta\%$	$\bar{T}$	$\mu_T$	$\Delta\%$
10	5	273.17	275.67	-0.91	136.58	137.76	-0.85	54.63	56.44	-3.20
40	5	2787.28	2821.84	-1.22	1393.64	1433.86	-2.80	557.46	598.03	-6.78
40	10	1176.20	1197.19	-1.75	588.10	604.54	-2.72	235.24	248.15	-5.20
80	5	9939.48	9888.86	+0.51	4969.74	5181.18	-4.08	1987.90	2164.12	-8.14
80	20	1598.81	1617.58	-1.16	799.41	824.12	-3.00	319.76	342.19	-6.55
100	5	15151.93	15183.58	-0.21	7575.97	7807.49	-2.97	3030.39	3267.66	-7.26
100	25	1763.36	1768.67	-0.30	881.68	903.18	-2.38	352.67	369.58	-4.58

$N$	$n$	$\rho = 0.10$			$\rho = 0.50$		
		$\bar{T}$	$\mu_T$	$\Delta\%$	$\bar{T}$	$\mu_T$	$\Delta\%$
10	5	27.32	29.47	-7.31	5.46	8.72	-37.35
40	5	278.73	319.18	-12.67	55.75	124.99	-55.40
40	10	117.62	132.71	-11.37	23.52	45.54	-48.35
80	5	993.95	1159.32	-14.26	198.79	495.32	-59.87
80	20	159.88	182.86	-12.57	31.98	62.29	-48.66
100	5	1515.19	1780.02	-14.88	303.04	769.96	-60.64
100	25	176.34	198.62	-11.22	35.27	70.41	-49.91

TABLE 1. Average times in the real settings and the corresponding low-traffic approximation for different values of  $N$ ,  $n$  and  $\rho$ .

under the low-traffic approximation  $\mu_T$  are given for several settings. For  $\rho$  up to 0.05 the approximation works well, and the relative error is less than 10% in all settings, while for  $\rho = 0.1$  and  $\rho = 0.5$  the differences become relevant, especially in the latter case.

#### 4. SUMMARY AND CONCLUSIONS

In this paper, we characterized the ergodic regime of a simple model for the continuous double auction in the low-traffic limit  $\rho \ll 1$ . In this limit, the price distribution can be derived for any value of the model parameters  $n$  and  $N$ . Explicit numerical procedures to find the price distributions are given in the Supplemental Material. We also showed that these results give a reasonable approximation of the auction behaviour for  $\rho < 1/2$ . We further studied the first passage time  $T$  in 1 or  $N$  using Monte Carlo simulations. We noticed that the low-traffic limit approximation for  $T$  works reasonably for  $\rho \ll 1/2$  in this case.

There are several open questions we would like to answer. A natural extension of this simple model is its semi-Markov version in which non exponential distributions for waiting times between events are introduced. In such an extension, the behavior of the embedded chain does not change, but the mixing time of the chain changes.

A particularly interesting case is when the distribution of waiting times is heavy-tailed with infinite mean. This is linked to recent results of ours on semi-Markov graph dynamics [19, 20]. A further research direction worth exploring is considering non-independent processes for limit and market orders.

#### ACKNOWLEDGEMENTS

T. Radivojević and E. Scalas wish to thank J. Anselmi for useful discussion. E. Scalas thanks N. Georgiou for discussion on multi-class queues.

#### REFERENCES

- [1] J. Blanchet and X. Chen, Continuous-time modeling of bid-ask spread and price dynamics in limit order books, arXiv:1310.1103 [q-fin.TR] (2013).
- [2] P. A. Ferrari, J. B. Martin, Multiclass Hammersley-Aldous-Diaconis process and multiclass-customer queues, *Ann. Inst. H. Poincaré*, **45**, 250–265 (2009).
- [3] T. Radivojević, J. Anselmi and E. Scalas, Ergodic Transition in a Simple Model of the Continuous Double Auction, *PLOS ONE* **9**, e88095 (2014).
- [4] E. Scalas, T. Kaizoji, M. Kirchler, J. Huber, and A. Tedeschi, Waiting times between orders and trades in double-auction markets, *Physica A* **366**, 463–471 (2006).
- [5] E. Smith, J. Doyne Farmer, L. Gillemot, and S. Krishnamurthy, Statistical theory of the continuous double auction, *Quantitative Finance* **3**, 481–514, (2002).
- [6] S. Zuo, M. Li, K. Kishimoto, M. Endo, A queuing model for a continuous double-auction trading system. In: *22nd Australasian Finance and Banking Conference* (2009).
- [7] D.K. Gode and S. Sunder, Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of Political Economy* **101**, 119–137 (1993).
- [8] T. Radivojević, J. Anselmi, and E. Scalas, A stylized model for the continuous double auction. In: *Managing Market Complexity*, Lecture Notes in Economics and Mathematical Systems **662**, 115–125, Springer, Berlin (2012).
- [9] R. Cont, A. de Larrard, Price dynamics in a Markovian limit order market, *SIAM Journal on Financial Mathematics* **4**, 1–25 (2013).
- [10] R. Cont, A. de Larrard, Order book dynamics in liquid markets: Limit theorems and diffusion approximations, Working paper available at SSRN: <http://ssrn.com/abstract=1757861> (2012).
- [11] J.F.C. Kingman, The single server queue in heavy traffic, *Mathematical Proceedings of the Cambridge Philosophical Society* **57**, 902–904 (1961). J.F.C. Kingman, On queues in heavy traffic, *Journal of the Royal Statistical Society Series B* **24**, 383–392 (1962).
- [12] W. Witt, *Stochastic-Process Limits*, Springer, New York (2002).
- [13] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, URL: <http://www.R-project.org/> (2014).
- [14] L. Bondesson, A characterization of first passage time distributions for random walks, *Stochastic Processes and Their Applications* **39**, 81–88 (1991).
- [15] T. Antal and S. Redner, Escape of a Uniform Random Walk from an Interval, *Journal of Statistical Physics* **123**, 1129–1144 (2006).
- [16] W. Feller, *An Introduction to Probability Theory and Its Applications. Volume I*, John Wiley & Sons, Inc., New York-London-Sydney (1968).

- [17] A. Carlsund, First passage time for random walks and birth-and-death processes with sign depending transition probabilities, Working Paper, Royal Institute of Technology, Stockholm (2000).
- [18] J.G. Kemeny and J.L. Snell, *Finite Markov chains*, Springer-Verlag, New York (1976).
- [19] M. Raberto, F. Rapallo and E. Scalas, Semi-Markov graph dynamics, PLOS ONE **6**, e23370 (2011).
- [20] N. Georgiou, I. Kiss and E. Scalas, Solvable non-Markovian dynamic network, Phys. Rev. E **92**, 042801 (2015).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SUSSEX, BRIGHTON, UK AND BCAM,  
BASQUE CENTER FOR APPLIED MATHEMATICS, BILBAO, SPAIN

*E-mail address:* e.scalas@sussex.ac.uk

DISIT, UNIVERSITÀ DEL PIEMONTE ORIENTALE, ALESSANDRIA, ITALY

*E-mail address:* fabio.rapallo@uniupo.it

BCAM, BASQUE CENTER FOR APPLIED MATHEMATICS, BILBAO, SPAIN

*E-mail address:* tradivojevic@bcamath.org

## APPENDIX A

In this appendix, we present a general algorithm to find the invariant distribution for the price Markov chain. First, we observe that, in general, the transition probability matrix is a stochastic block tri-diagonal (symmetric in the inner part) matrix of the form:

$$(10) \quad P = \begin{pmatrix} D_0 & A & 0 & \dots & 0 \\ A^T & D & A & \dots & 0 \\ 0 & A^T & D & \dots & 0 \\ \vdots & & & & D & A \\ 0 & \dots & & A^T & \widetilde{D}_0 \end{pmatrix},$$

where  $A^T$  is the transpose of the block  $A$  and  $\widetilde{D}_0 = (d_{n-i+1, n-j+1})$  for  $d_{ij}$  being elements of the block  $D_0$ . There is some freedom in the choice of the blocks  $D_0$ ,  $D$  and  $A$  and their transformations whose dimensions depend on the value of  $n$ . In the case of the transition probability (4) in the paper, this is a possible choice of blocks:

$$(11) \quad P = \begin{pmatrix} 4/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 5/12 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 5/12 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 4/6 \end{pmatrix},$$

but the following block choice seems to be more convenient

$$(12) \quad P = \begin{pmatrix} 4/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/4 & 5/12 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/3 & 1/6 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 5/12 & 1/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 4/6 \end{pmatrix}.$$

To better see how this generalizes, let us consider the structure of the matrix for  $N = 10$  and  $n = 2$ , once more

$$(13) \quad P = \begin{pmatrix} d_1 & a & a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ b_2 & d_2 & a & a & 0 & 0 & 0 & 0 & 0 & 0 \\ a & a & 2a & a & a & 0 & 0 & 0 & 0 & 0 \\ 0 & a & a & 2a & a & a & 0 & 0 & 0 & 0 \\ 0 & 0 & a & a & 2a & a & a & 0 & 0 & 0 \\ 0 & 0 & 0 & a & a & 2a & a & a & 0 & 0 \\ 0 & 0 & 0 & 0 & a & a & 2a & a & a & 0 \\ 0 & 0 & 0 & 0 & 0 & a & a & 2a & a & a \\ 0 & 0 & 0 & 0 & 0 & 0 & a & a & d_2 & b_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & a & a & d_1 \end{pmatrix}.$$

For  $N = 10, n = 3$ , instead, we have

$$(14) \quad P = \begin{pmatrix} d_1 & a & a & a & 0 & 0 & 0 & 0 & 0 & 0 \\ b_2 & d_2 & a & a & a & 0 & 0 & 0 & 0 & 0 \\ b_3 & b_3 & d_3 & a & a & a & 0 & 0 & 0 & 0 \\ a & a & a & 2a & a & a & a & 0 & 0 & 0 \\ 0 & a & a & a & 2a & a & a & a & 0 & 0 \\ 0 & 0 & a & a & a & 2a & a & a & a & 0 \\ 0 & 0 & 0 & a & a & a & 2a & a & a & a \\ 0 & 0 & 0 & 0 & a & a & a & d_3 & b_3 & b_3 \\ 0 & 0 & 0 & 0 & 0 & a & a & a & d_2 & b_2 \\ 0 & 0 & 0 & 0 & 0 & 0 & a & a & a & d_1 \end{pmatrix}$$

and so on, where:

$$(15) \quad \begin{aligned} a &= \frac{1}{2(n+1)} \\ d_i &= \frac{1}{2i} + a \\ b_i &= \frac{1}{2i} \end{aligned}$$

as a consequence of (3) in the paper and

$$(16) \quad \begin{aligned} d_1 + na &= 1 \\ b_2 + d_2 + na &= 1 \\ b_3 + b_3 + d_3 + na &= 1 \\ &\vdots \\ (n-1)b_n + d_n + na &= 1 \\ 2na + 2a &= 1 \end{aligned}$$

as a consequence of the properties of the transition matrix.

The linear system of equations whose solution is the invariant distribution from Eq. (5) in the paper is:

$$(17) \quad \begin{cases} d_1\pi_1 + b_2\pi_2 + \cdots + b_n\pi_n + a\pi_{n+1} & = \pi_1 \\ a\pi_1 + d_2\pi_2 + b_3\pi_3 + \cdots + b_n\pi_n + a\pi_{n+1} + a\pi_{n+2} & = \pi_2 \\ \vdots & \vdots \\ a\pi_1 + \cdots + a\pi_{n-1} + d_n\pi_n + a\pi_{n+1} + \cdots + a\pi_{2n} & = \pi_n \\ a\pi_1 + \cdots + a\pi_n + 2a\pi_{n+1} + a\pi_{n+2} \cdots + a\pi_{2n+1} & = \pi_{n+1} \\ a\pi_2 + \cdots + a\pi_{n+1} + 2a\pi_{n+2} + a\pi_{n+3} \cdots + a\pi_{2n+2} & = \pi_{n+2} \\ \vdots & \vdots \\ a\pi_{N-2n} + \cdots + a\pi_{N-n-1} + 2a\pi_{N-n} + a\pi_{N-(n-1)} \cdots + a\pi_N & = \pi_{N-n} \\ a\pi_{N-2n-1} + \cdots + a\pi_{N-n} + d_n\pi_{N-(n-1)} + a\pi_{N-(n-2)} + \cdots + a\pi_N & = \pi_{N-(n-1)} \\ \vdots & \vdots \\ a\pi_{N-n} + b_n\pi_{N-(n-1)} \cdots + b_2\pi_{N-1} + d_1\pi_N & = \pi_N \end{cases}$$

with the additional equation:

$$(18) \quad \sum_{i=1}^N \pi_i = 1.$$

A possible numerical **solver** for the system (17) given  $N$  and  $n$ , and written in MATLAB is:

```
function prob = pricedistrlt(N,n)
a = 1/2/(n+1);
% defining the transition matrix
d = 2*a*ones(1,N); % blue diagonal values
d(1:n) = 0.5./[1:n] + a; % green diagonal values
P = diag(d);
for i=1:n % red
    P = P + diag(a*ones(1,N-i),i) + diag(a*ones(1,N-i),-i);
end
for i=2:n % green
    for j=1:i-1
        P(i,j) = 1/2/i;
    end
end
P(N-n+1:N,N-n+1:N) = rot90(P(1:n,1:n),2);

X = sym('x',[1 N]);
x = solve(X*P - X,sum(X)-1);
x = struct2cell(x);
```



```
prob = zeros(1,N);  
for i=1:N  
    prob(i) = x{i};  
end
```

The above function requires the Symbolic Math Toolbox in MATLAB. Another possibility is to solve (17) using the Matlab function `eig` after defining the transition matrix in the following way:

```
[prob,1] = eig(P. ');  
if n==1  
    prob = prob(:,end)./sum(prob(:,end));  
else  
    prob = prob(:,1)./sum(prob(:,1));  
end
```

This option is much faster, but less accurate.