

HYDRODYNAMIC LIMIT OF ORDER-BOOK DYNAMICS

XUEFENG GAO

*Department of Systems Engineering and Engineering Management,
The Chinese University of Hong Kong,
Shatin, N.T.
Hong Kong
E-mail: xfgao@se.cuhk.edu.hk*

S.J. DENG

*H. Milton Stewart School of Industrial and Systems Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332,
USA
E-mail: deng@isye.gatech.edu*

In this paper, we establish a fluid limit for a two-sided Markov order book model. The main result states that in a certain asymptotic regime, a pair of measure-valued processes representing the “sell-side shape” and “buy-side shape” of an order book converges to a pair of deterministic measure-valued processes in a certain sense. We also test the fluid approximation on data. The empirical results suggest that the approximation is reasonably good for liquidly traded stocks in certain time periods.

Keywords: applied probability, mathematical finance, queueing theory, stochastic modeling

1. INTRODUCTION

As a trading mechanism, limit order books have gained growing popularity in equity and derivative markets in the past two decades. Nowadays, the majority of the world’s financial markets, such as Electronic Communication Networks in the USA, the Hong Kong Stock Exchange and the Toronto Stock Exchange, are organized as electronic limit order books to match buyers and sellers. This has inspired intensive research activities on limit order books. See, for example, [12,28] for reviews.

The primary goal of this paper is to develop approximations for the evolution of the shape of a limit order book over a time horizon that is large compared with the length of time between order book events. To this end, we perform asymptotic analysis of a stochastic limit order-book model and thus connect the dynamics of a limit order book on two different time scales. There has been a growing interest in studying such scaling limits of order book dynamics, mainly motivated by a desire to better understand the price formation process and the relation between the parameters of the point processes describing order flow at high

frequency and the parameters of models describing price and order-book shape dynamics at a larger time scale. See, for example, [1,5,7,8,15,19,24,25].

We contribute to this body of literature by establishing a fluid limit for a two-sided in the following asymptotic regime: (a) Tick size goes to zero; (b) Rates of order arrivals goes to infinity; (c) Relative order size (compared with limit order queue size) goes to zero. Such a regime is relevant for high-frequency trading of liquid stocks where the tick size is small (one penny in the USA), inter-arrival times of orders are very short (on the order of milliseconds), and there are large volumes of limit orders sitting near the market price and waiting for transactions.

Besides establishing a scaling limit, we also test the fluid approximation on historical order book data from the U.S. exchange NYSE Arca. The in-sample analysis demonstrates that our theory can be potentially useful to approximate the evolution of a limit order book for liquid stocks in time periods of low-price volatility.

Our asymptotic analysis is built on the order-book model in [9], where the high-frequency dynamics of order book events are described by a continuous-time Markov chain (CTMC). In their discrete order-book model, there are a finite number of security price levels and the state of the order book is described by an order quantity at each price. We develop a fluid approximation for this high-dimensional CTMC by establishing a type of law of large numbers limit theorem.

The main result (Theorem 3.1) states that as the tick size $1/n$ approaches zero (the tick size, the arrival rates of orders, and the relative order sizes are all scaled as a power of a large parameter n . See Section 2 for more details.), (a) the sequence of best bid and best ask prices converges in probability to a constant; and (b) the sequence of pairs of measure-valued processes $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$, representing the “sell-side shape” and “buy-side shape” of the order book, converges to a pair of deterministic measure-valued processes in a certain sense. Moreover, the density profile of the limiting processes can be described by Ordinary Differential Equations (ODEs) with coefficients determined by order flow intensities.

Our key innovation in establishing the fluid limit is to represent a *two-sided limit order book* by a *pair of non-negative measure-valued processes* and studying the weak convergence of such processes using martingale methods. Such a measure-valued process representation of an order-book and the related proof techniques could be potentially useful for establishing scaling limits for more general two-sided order-book models such as multi-scale models allowing the co-existence of high-frequency and low-frequency traders. See, for example, [6].

We now compare our work with some closely related alternatives which also study the scaling limits of the shape of limit order books in certain asymptotic regimes. These include [3,14,15,21,22,24,25]. Among them, only the three studies [3,14,15] incorporate order cancellations. Modelling order cancellations is important: e.g., more than 95% of limit orders at NASDAQ are canceled without execution, see [13]. In [15], the authors prove a law of large numbers result for the limit order book dynamics. In their limit, the best bid and best ask price dynamics can be described by two coupled ODEs and the relative buy and sell shape functions can be described by two linear first-order partial differential equations. This fluid limit result is later generalized in [14] to allow order flows depend on the state of the order book including the prices and volumes, and in [3] which establishes the functional central limit theorems for order book dynamics. We remark that though we rely on a slightly different discrete stochastic order-book model to conduct asymptotic analysis, our theoretical result is not completely new in view of these three studies. However, we emphasize that our proof technique to establish the scaling limit is novel and it is very different from theirs. Moreover, unlike the aforementioned studies, we empirically test our fluid approximation on real data and illustrate the potential relevance of the approximation. These are the two main contributions of this paper.

Finally, we remark that we have had to focus on a few features to keep the model and the asymptotic analysis tractable, and that this has necessarily led to some undesirable consequences. We briefly discuss the main limitations. First, the limit is deterministic, and the limit price is a constant. This is not realistic. However, such a deterministic model could be potentially useful. For example, a similar fluid order-book model, where the best bid and ask prices are also constants, has been used in [26] to analyze the problem of limit and market order placement to optimally buy a block of shares over a fixed time horizon in the order of several minutes. Second, to keep the mathematical analysis tractable, a number of realistic features such as non-stationarity, time-clustering, and mutual dependence of order flows are necessarily left out in our asymptotic analysis. For recent developments along this line, see, for example, [1].

Outline of this paper. Section 2 reviews a variant of the Markov order-book model in [9] and states the assumptions on order flow rates and initial conditions. Section 3 summarizes our main result. Section 4 discusses empirical analysis. Section 5 presents preliminaries for proving the main result. Section 6 is devoted to the proof of the main result. Auxiliary results and their proofs are provided in the appendices.

Notation. All random elements are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ unless otherwise specified. Given $x \in \mathbb{R}$, we set $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$. We write $\lceil x \rceil$ for the smallest integer not less than x , and $\lfloor x \rfloor$ for the largest integer not greater than x . The set of continuous functions on $[0, 1]$ is denoted by $C([0, 1])$. Given a Polish space \mathcal{E} , the space of right-continuous functions $f : [0, T] \rightarrow \mathcal{E}$ with left limits is denoted by $\mathbb{D}([0, T], \mathcal{E})$. The space $\mathbb{D}([0, T], \mathcal{E})$ is assumed to be endowed with the Skorohod J_1 -topology. For a sequence of random elements $\{X_n : n \geq 1\}$ taking values in a metric space, we write $X_n \Rightarrow X$ to denote the convergence of X_n to X in distribution. Each stochastic process with sample paths in $\mathbb{D}([0, T], \mathcal{E})$ is considered to be a $\mathbb{D}([0, T], \mathcal{E})$ -valued random element. For a Borel measure ν and function f , we set $\langle \nu, f \rangle = \int f(u)\nu(du)$ when the integration exists. The symbol δ_u represents the Dirac measure at location $u \in \mathbb{R}$, that is, for a Borel set U ,

$$\delta_u(U) = \begin{cases} 1 & \text{if } u \in U, \\ 0 & \text{if } u \notin U. \end{cases}$$

The space of finite *non-negative* measures on $[0, 1]$ is denoted by $\mathcal{M}^+([0, 1])$, and the space of finite *signed* measures on $[0, 1]$ is denoted by $\mathcal{M}([0, 1])$.

2. MODEL AND ASSUMPTIONS

In this section we describe a variant of the order-book model introduced in [9] and state the assumptions. Throughout of this paper, we fix the time $T > 0$.

Fix $n \geq 1$. Without loss of generality, we suppose that investors can submit their limit orders to n discrete equally spaced price levels $\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n}{n}\}$ within the price range $(0, 1]$. Thus, the parameter n also represents the inverse of tick size. The state of the limit order book at time t is given by a vector $\mathcal{X}^n(t) \equiv (\mathcal{X}_1^n(t), \dots, \mathcal{X}_n^n(t)) \in \mathbb{Z}^n$, where for $i \in \{1, \dots, n\}$, $|\mathcal{X}_i^n(t)|$ represents the number of outstanding limit orders at price i/n at time t . If $\mathcal{X}_i^n(t) > 0$, then there are $\mathcal{X}_i^n(t)$ sell orders at price i/n , and if $\mathcal{X}_i^n(t) < 0$, then there are $-\mathcal{X}_i^n(t)$ buy orders at price i/n .

To define the best ask price (lowest price among limit sell orders) and best bid price (highest price among limit buy orders), we define two mappings \mathcal{P}_A^n and \mathcal{P}_B^n such that for

a given state $x = (x_1, \dots, x_n) \in \mathbb{Z}^n$

$$\mathcal{P}_A^n(x) \equiv \inf\{i \in \{1, \dots, n\} : x_i > 0\} \wedge \{n + 1\}, \tag{2.1}$$

$$\mathcal{P}_B^n(x) \equiv \sup\{i \in \{1, \dots, n\} : x_i < 0\} \vee 0, \tag{2.2}$$

where $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$ by convention. For $t \geq 0$, we write

$$p_A^n(t) = \mathcal{P}_A^n(\mathcal{X}^n(t)), \tag{2.3}$$

$$p_B^n(t) = \mathcal{P}_B^n(\mathcal{X}^n(t)). \tag{2.4}$$

Thus for order book $\mathcal{X}^n(\cdot)$ at time t , the best ask price is $p_A^n(t)/n$ and best bid price is $p_B^n(t)/n$, where $1/n$ is the tick size. We will be interested in the regime where $n \rightarrow \infty$, that is, the tick size approaches zero. For each fixed n , as in [9], we assume

- (a) Limit buy (respectively sell) orders arrive at a distance of i ticks from the opposite best quote at independent, exponentially distributed times with rate $\Lambda_B^n(i)$ (respectively $\Lambda_A^n(i)$),
- (b) Market buy (respectively sell) orders arrive at independent, exponentially distributed times with rate Υ_B^n (respectively Υ_A^n),
- (c) Each limit buy (respectively sell) order at a distance of i ticks from the opposite best quote is canceled independently after exponentially distributed times with rate $\Theta_B^n(i)$ (respectively $\Theta_A^n(i)$).
- (d) All the above events are mutually independent.
- (e) All the orders are of unit size, which is independent of n .

Compared to the model in [9], the extra feature is that we allow order flow intensities to be side-dependent (buy orders or sell orders). This has been observed empirically and such feature will be useful in our empirical analysis in Section 4. Given these assumptions, the state process $\mathcal{X}^n(\cdot)$ is a n -dimensional CTMC with state space \mathbb{Z}^n . Define an operator \mathcal{L}_n as follows: for any function $\mathcal{H} : \mathbb{Z}^n \rightarrow \mathbb{R}$ and each $x \in \mathbb{Z}^n$

$$\begin{aligned} \mathcal{L}_n \mathcal{H}(x) = & \sum_{k < \mathcal{P}_A^n(x)} [(\mathcal{H}(x^{k+}) - \mathcal{H}(x)) \cdot \Theta_B^n(\mathcal{P}_A^n(x) - k) \cdot |x_k|] \\ & + \sum_{k < \mathcal{P}_A^n(x)} [(\mathcal{H}(x^{k-}) - \mathcal{H}(x)) \cdot \Lambda_B^n(\mathcal{P}_A^n(x) - k)] \\ & + \sum_{k > \mathcal{P}_B^n(x)} [(\mathcal{H}(x^{k+}) - \mathcal{H}(x)) \cdot \Lambda_A^n(k - \mathcal{P}_B^n(x))] \\ & + \sum_{k > \mathcal{P}_B^n(x)} [(\mathcal{H}(x^{k-}) - \mathcal{H}(x)) \cdot \Theta_A^n(k - \mathcal{P}_B^n(x)) \cdot |x_k|] \\ & + (\mathcal{H}(x^{\mathcal{P}_B^n(x)+}) - \mathcal{H}(x)) \cdot \Upsilon_A^n \\ & + (\mathcal{H}(x^{\mathcal{P}_A^n(x)-}) - \mathcal{H}(x)) \cdot \Upsilon_B^n, \end{aligned} \tag{2.5}$$

where \mathcal{P}_A^n and \mathcal{P}_B^n are mappings given in (2.1) and (2.2), and for $k \in \{1, \dots, n\}$

$$\begin{cases} x^{k+} = (x_1, \dots, x_{k-1}, x_k + 1, x_{k+1}, \dots, x_n), \\ x^{k-} = (x_1, \dots, x_{k-1}, x_k - 1, x_{k+1}, \dots, x_n), \end{cases}$$

and

$$x^{0+} = x^{(n+1)-} = x.$$

The infinitesimal generator of \mathcal{X}^n coincides with the operator \mathcal{L}_n on the space of bounded functions from \mathbb{Z}^n to \mathbb{R} .

A few remarks are in order. First, the price range $(0, 1]$ is not essential in our asymptotic analysis. We can also use $(c_1, c_2]$ for some integers $c_1 < c_2$ that are independent of n , and this would not affect our results. Second, the state space of the above model is the n -dimensional integer lattice and it becomes infinitely dimensional as $n \rightarrow \infty$. So standard limit theorems on Markov chains (see, e.g., [23]) do not apply in our asymptotic analysis. Third, it is possible to allow random order size with finite second moment in our analysis. For example, if we adopt the order-book model in [16] which generalizes [9] by allowing limit orders to arrive in multiple sizes, we still have the Markovian structure for the order-book dynamics and our main result on the fluid limit would essentially not change. To keep the presentation simple, as in [9], we work with unit order size (which can be taken as the average order size) in this paper.

Throughout the paper, we make the following assumptions.

ASSUMPTION 2.1: *There exists a function ϱ from $[0, 1]$ to \mathbb{R} such that for fixed $n \geq 1$, the initial order book $\mathcal{X}^n(0)$ is given by*

$$\mathcal{X}_i^n(0) \equiv n \cdot \varrho\left(\frac{i}{n}\right) \quad \text{for } i \in \{1, 2, \dots, n\}. \tag{2.6}$$

The initial profile ϱ satisfies

(a) For some $\mathbf{p} \in (0, 1)$, $\varrho(\mathbf{p}) = 0$ and

$$\sup\{x \in [0, 1] : \varrho(x) < 0\} = \inf\{x \in [0, 1] : \varrho(x) > 0\} = \mathbf{p}; \tag{2.7}$$

(b) the function $|\varrho|$ is positive on $(\mathbf{p} - \Delta, \mathbf{p})$ and $(\mathbf{p}, \mathbf{p} + \Delta)$ for some small $\Delta > 0$;

(c) the function ϱ is continuous and bounded on $[0, \mathbf{p})$ and $(\mathbf{p}, 1]$.

ASSUMPTION 2.2: *For each market side $j \in \{A, B\}$, there exists positive continuous function $\Lambda_j(x)$ and non-negative continuous function $\Theta_j(x)$ on $[0, 1]$, and constants $\kappa < 1$ and $\Upsilon \geq 0$ such that for each n and $i \in \{1, 2, \dots, n\}$,*

$$\Lambda_j^n(i) = \Lambda_j\left(\frac{i}{n}\right), \quad \Theta_j^n(i) = \frac{1}{n}\Theta_j\left(\frac{i}{n}\right), \quad \text{and} \quad \Upsilon_j^n \leq n^\kappa \cdot \Upsilon. \tag{2.8}$$

Both assumptions are motivated by our data analysis. First, empirically we find that for certain liquidly traded stocks such as Bank of America, Ford Motor Co. and Dell Inc., there are large volumes of limit orders sitting at or near the best quotes and waiting for executions. This provides empirical support for (2.6). For technical convenience, we also make three assumptions on the regularity of the initial profile ϱ . Here, the number \mathbf{p} can be viewed as the “market price” of a stock at time zero. Second, from our data we also find that for the aforementioned liquid stocks, typically the ratio of limit order arrival rate and cancel rate $\frac{\Lambda_j^n(i)}{\Theta_j^n(i)}$ is of the order of $n = 100$ in the vicinity of the best quotes. In addition, we observe that the magnitude of market order arrival rate (e.g., volume per second) is not very high compared to the volumes of limit orders in the vicinity of market price, specifically during

the time periods of low price volatility. These two facts motivate (2.8). The assumptions on the continuity of functions Λ_j, Θ_j and positivity of Λ_j are technical and they facilitate our mathematical analysis.

3. THE MAIN RESULT

In this section, we state the main result of this paper. We first define three sequences of measure-valued process (see, e.g., [10] for background on measure-valued processes). For fixed $n \geq 1$ and $t \in [0, T]$, we set

$$\zeta_t^{n,+} = \frac{1}{n^2} \sum_{i=1}^n (\mathcal{X}_i^n(nt))^+ \cdot \delta_{\frac{i}{n}} = \frac{1}{n^2} \sum_{i > p_B^n(nt)} \mathcal{X}_i^n(nt) \cdot \delta_{\frac{i}{n}}, \tag{3.1}$$

$$\zeta_t^{n,-} = \frac{1}{n^2} \sum_{i=1}^n (\mathcal{X}_i^n(nt))^- \cdot \delta_{\frac{i}{n}} = \frac{1}{n^2} \sum_{i < p_A^n(nt)} -\mathcal{X}_i^n(nt) \cdot \delta_{\frac{i}{n}}, \tag{3.2}$$

$$\zeta_t^n = \zeta_t^{n,+} - \zeta_t^{n,-} = \frac{1}{n^2} \sum_{i=1}^n \mathcal{X}_i^n(nt) \cdot \delta_{\frac{i}{n}}, \tag{3.3}$$

where \mathcal{X}^n is the n -dimensional CTMC with state space \mathbb{Z}^n and it describes the evolution of the limit order book, p_A^n, p_B^n are given in (2.3) and (2.4), and δ_u is the Dirac measure centered at u . In (3.1)–(3.3), we scale the time by a large number n so that it allows us to study the low-frequency dynamics of an order book based on the high-frequency dynamical model \mathcal{X}^n . We scale the space by n^2 so that the pair $(\zeta^{n,+}, \zeta^{n,-})$ will have a non-trivial limit as $n \rightarrow \infty$. Let us give an intuitive explanation on this combination of space-time scalings. When we scale the time and speed up the order-book process \mathcal{X}^n by a factor of n , the rates of order arrivals are then magnified from order one to order n . Together with Assumption 2.1 which indicates the initial queue size of limit orders is of the order n , one can then expect that the number of limit orders on each price level is of the order n at each time t . Hence, the choice of space scaling n^2 follows since there are n price levels in total and the total number of sell limit orders and buy limit orders at all n price levels are of the order of n^2 .

For fixed t , the pair $(\zeta_t^{n,+}, \zeta_t^{n,-})$ represents the “sell-side shape” and “buy-side shape” of the order book at some large time nt , and the measure ζ_t^n represents the whole shape of the order book at large time nt . Note that the pair $(\zeta_t^{n,+}, \zeta_t^{n,-})$ lives in the product space $\mathcal{M}^+([0, 1]) \times \mathcal{M}^+([0, 1])$. For the moment, we use $\overline{\mathcal{M}}([0, 1])$ to denote this product space equipped with an appropriate metric such that $\overline{\mathcal{M}}([0, 1])$ is complete and separable. A precise definition of the space $\overline{\mathcal{M}}([0, 1])$ is given in Section 5.2.

We are interested in the limiting behavior of the Markov process $(\zeta^{n,+}, \zeta^{n,-})$ as $n \rightarrow \infty$. This corresponds to the limiting regime where price tick size $1/n$ approaches zero, rates of order arrivals goes to infinity, and relative order size (compared to limit order queue size) goes to zero. This can be readily seen from Assumptions 2.1–2.2 and the fact that the time is speeded up by n in (3.1)–(3.2). It is worthwhile to note that ζ^n is also a Markov process, which contains the same information as $(\zeta^{n,+}, \zeta^{n,-})$. However, for fixed t , ζ_t^n is a signed measure living in the space $\mathcal{M}([0, 1])$ which is ill-suited for studying weak convergence since the weak topology is in general not metrizable (see, e.g., [31]). Thus we instead work with the sequence of pairs of non-negative measure-valued processes $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$.

Now we state the main theorem of this paper. The proof is given in Section 6. Recall that $\varrho, p, \Lambda_A, \Lambda_B, \Theta_A$, and Θ_B are given in Assumptions 2.1 and 2.2.

THEOREM 3.1: *Suppose that Assumptions 2.1 and 2.2 hold. Then as $n \rightarrow \infty$,*

(a) *we have*

$$\sup_{0 \leq t \leq T} \left| \frac{p_A^n(nt)}{n} - p \right| \Rightarrow 0, \tag{3.4}$$

$$\sup_{0 \leq t \leq T} \left| \frac{p_B^n(nt)}{n} - p \right| \Rightarrow 0; \tag{3.5}$$

(b) *we have*

$$(\zeta^{n,+}, \zeta^{n,-}) \Rightarrow (\zeta^+, \zeta^-) \text{ in } \mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1])),$$

where (ζ^+, ζ^-) is a pair of deterministic measure-valued process. In addition, for any $t \in [0, T]$, the non-negative measures ζ_t^+ and ζ_t^- are absolutely continuous with respect to Lebesgue measure and have density functions $\varphi^+ = \max\{\varphi, 0\}$ and $\varphi^- = \max\{-\varphi, 0\}$ such that $\zeta_t^+(dx) = \varphi^+(x, t)dx$ and $\zeta_t^-(dx) = \varphi^-(x, t)dx$ for $x \in [0, 1]$. The function φ is uniquely determined by the following set of equations: for $x \in [0, 1]$

$$\varphi(x, 0) = \varrho(x), \tag{3.6}$$

$$\partial_t \varphi(x, t) = \Lambda_A(x - p) - \Theta_A(x - p) \cdot \varphi(x, t), \quad x > p, \tag{3.7}$$

$$\partial_t \varphi(x, t) = -\Lambda_B(p - x) - \Theta_B(p - x) \cdot \varphi(x, t), \quad x < p. \tag{3.8}$$

$$\varphi(p, t) = 0, \quad t \in [0, T]. \tag{3.9}$$

Theorem 3.1 provides a fluid approximation for the order-book dynamics. If we allow limit orders to arrive in multiple sizes while keeping market orders and cancellations of existing orders being unit size as in [16] which generalizes [9], our fluid limit still holds and the only change we need to make is to multiply the limit order arrival rate Λ_A, Λ_B (number of orders per unit time) by the average limit order size (relative to the market order size) in Eqs. (3.7) and (3.8). Note that in high-frequency trading, the inter-arrival times of order book events are typically very short (in the order of milliseconds). So the result suggests that on a “low-frequency” time scale (in the order of a few minutes), the scaled best bid and best ask prices are a constant p , and the density profile of the sell-side and buy-side shape of the order book, given by $\varphi^+(x, t)$ and $\varphi^-(x, t)$, are characterized by a linear ODE with coefficients determined by order arrival rates and cancellation rates. As discussed in the introduction, this deterministic limit may not be realistic, but it could be potentially useful for studying other problems.

One readily verifies from Theorem 3.1 and (2.7) that for each $t \in [0, T]$,

$$\varphi^+(x, t) = e^{-\Theta_A(x-p)t} \cdot \varrho(x)^+ + \frac{\Lambda_A(x-p)}{\Theta_A(x-p)} \cdot \left(1 - e^{-\Theta_A(x-p)t}\right) \text{ for } p < x \leq 1, \tag{3.10}$$

$$\varphi^-(x, t) = e^{-\Theta_B(p-x)t} \cdot \varrho(x)^- + \frac{\Lambda_B(p-x)}{\Theta_B(p-x)} \cdot \left(1 - e^{-\Theta_B(p-x)t}\right) \text{ for } 0 \leq x < p. \tag{3.11}$$

When n is large, we obtain from Theorem 3.1 the following approximation for the transient behavior of the shape of an order book: for interval $I \subset [0, 1]$ and $t \in [0, T]$

$$\zeta_t^{n,+}(I) \triangleq \frac{1}{n^2} \sum_{i > p_B^n(nt), \frac{i}{n} \in I} \mathcal{X}_i^n(nt) \approx \zeta_t^+(I) = \int_I \varphi^+(x, t) dx \approx \frac{1}{n} \sum_{\frac{i}{n} \in I} \varphi^+\left(\frac{i}{n}, t\right), \tag{3.12}$$

$$\zeta_t^{n,-}(I) \triangleq \frac{1}{n^2} \sum_{i < p_A^n(nt), \frac{i}{n} \in I} -\mathcal{X}_i^n(nt) \approx \zeta_t^-(I) = \int_I \varphi^-(x, t) dx \approx \frac{1}{n} \sum_{\frac{i}{n} \in I} \varphi^-\left(\frac{i}{n}, t\right), \tag{3.13}$$

where φ^- and φ^+ are given in (3.10) and (3.11). In the next section, we will test the approximations (3.12) and (3.13) on data.

4. EMPIRICAL ANALYSIS

In this section, we perform in-sample analysis to illustrate the potential relevance of our fluid approximation in Theorem 3.1.

Our empirical analysis is based on one-month message-level order book data from NYSE Arca (in short, Arca) in August 2010. As of 2009, around 20% of market share for NASDAQ-listed securities and 10% of NYSE-listed securities are traded on Arca.

Two data sets are used. The first data set consists of all limit order activities on Arca, including limit order submission, modification and deletion. For each message in the data, it contains a time stamp down to millisecond, the price and order size, the buy or sell indicator, stock symbol, exchange, and an ID (identifier). This data set enables us to recreate the limit order book at any give time for stocks traded on Arca. Our second data set records all the trades. Each message contains a time stamp down to second, the traded price and order size, the buy or sell indicator, the best bid/ask prices and standing limit order quantities on these two prices when trades occur, stock symbol, and an ID (identifier). In conjunction with the first data set, we can analyze the detailed order flow properties such as limit order arrival rate, market order arrival rate and limit order cancellation rate.

4.1. Empirical Analysis

The goal of this section is to illustrate that our theoretical model can be potentially used to approximate the evolution of the shape of an order book on time scale of minutes for stocks and time periods, which satisfy our model assumptions.

We specifically test the approximations in (3.12) and (3.13) with empirical data. For convenience we multiply both sides of these approximate equations by n^2 , that is, we test the following approximations and show that they hold reasonably well:

$$\sum_{i > p_B^n(nt), \frac{i}{n} \in I} \mathcal{X}_i^n(nt) \approx n \sum_{\frac{i}{n} \in I} \varphi^+\left(\frac{i}{n}, t\right), \tag{4.1}$$

$$\sum_{i < p_A^n(nt), \frac{i}{n} \in I} -\mathcal{X}_i^n(nt) \approx n \sum_{\frac{i}{n} \in I} \varphi^-\left(\frac{i}{n}, t\right). \tag{4.2}$$

The left-hand side of (4.1) (respectively (4.2)) represents the total number of limit sell (respectively buy) orders in price interval $I \subset [0, 1]$. With an appropriate shift of the price levels, such quantities can be directly obtained from data since we have the full empirical order book information at each time instant. For example, we can simply shift the best ask

price to 0.5 and move other prices and order queues correspondingly. On the other hand, the quantities at the right-hand side involve the values of functions φ^+ and φ^- at discrete price levels i/n . We next discuss how to compute such values.

First, we screen the data to identify stocks and time periods, which satisfy Assumptions 2.1 and 2.2. We find that liquidly-traded stocks, such as Bank of America Corp. (BAC), General Electric Co. (GE), Ford Motor Co. (F) and Dell Inc. (DELL), typically have large volumes of limit orders sitting at or near the best quotes and waiting for market transactions. This is consistent with Assumption 2.1. Once a stock is chosen, we scan the intra-day order book data and look for time horizon (e.g., a few minutes) with low price volatility such that Assumption 2.2 is also satisfied. To give an intuitive explanation, note that when the volatility is low, trading is typically not active so the market order arrival rate is not large which is the key in Assumption 2.2. On the other hand, for the liquid stocks that we study, there are long queues of limit orders sitting in the order book. So if the market order arrival rate is not relatively large in certain periods, it will take long time for the quote to be depleted and for the price to move. This leads to the low volatility of the stocks in such periods. For illustration purposes, we focus on the stock BAC during 12:45–13:05 pm on August 5, 2010 as a representative example. The best bid and ask prices of BAC barely change during this 20 min time period.

Second, we estimate and input the needed parameters to compute φ^+ and φ^- given in (3.10) and (3.11). Such parameters include the initial order book profile ρ , the number \mathfrak{p} , the time length nt , the limit order arrival rates and the limit order cancellation rates. We follow the four-step procedure described below to choose and compute these parameters.

1. Normalize the calendar time 12:45 pm to be time zero. The initial order book is given by the first snapshot of the order book after 12:45 pm. Hence, we obtain ρ at each discrete price level.
2. Set \mathfrak{p} to be the best bid price at 12:45 pm when computing φ^+ and the best ask when computing φ^- . Best bid and ask prices at 12:45 pm are \$13.99 and \$14, respectively.
3. Take snapshots of the order book every 5 min till 13:05 pm. More precisely, we record the first snapshot of the order book after 12:50, 12:55, 13, and 13:05 pm, respectively. We do so because we are interested in the dynamics of the shape on the “large” time scale in the order of minutes.
4. For each of the four time intervals: 12:45–12:50 pm, 12:45–12:55 pm, 12:45–13:00 pm and 12:45–13:05 pm, estimate the order flow intensities following [9].

Finally, we compute the right-hand side of (4.1) and (4.2) by plugging in the model parameters. We use $n = 100$ which is the reciprocal of the tick size one cent. For the four time points 12:50, 12:55, 13:00, and 13:05 pm, we have nt correspond to 5, 10, 15, and 20 min, respectively. We take interval I in (4.1) and (4.2) as 3-cent price bins, e.g., [\$0.5, \$0.53]. It turns out that varying the bin size yields similar results, so we suppress further details.

We now present the main empirical findings on in-sample tests of the approximations in (4.1) and (4.2). We use Table 1 and Figure 1 for an illustration. For further statistics and results on both buy-side and sell-side shapes, see the Internet supplement available at the author’s website <http://www1.se.cuhk.edu.hk/xfgao/lobSupplement>. For the clarity of the presentation, we focus on price levels 10 ticks above the best ask price. Table 1 records the values of the function $\varphi^+(x, t)$ for stock BAC at five different time points including 12:45 pm, which corresponds to the time zero. From Table 1 and (4.1), one can readily obtain Figure 1, which plots the comparisons of the empirical sell-side shape of stock BAC and the theoretical sell-side shape obtained from applying our model at four different time instants.

TABLE 1. This table records the values of $\varphi^+(x, t)$ for BAC during 12:45–13:05 pm, August 5, 2010. The time $t = 0$ (unit, minute) corresponds to 12:45 pm. The position $x = \$0.50$ corresponds to \$14.00, the best ask price at 12:45 pm due to the price shift. When computing φ^+ , the number p is set to be \$0.49 which corresponds to the best bid price \$13.99.

| x | t | | | | |
|------|-------|-------|-------|-------|-------|
| | 0 | 0.05 | 0.1 | 0.15 | 0.2 |
| 0.50 | 1.355 | 1.845 | 2.265 | 2.780 | 2.998 |
| 0.51 | 1.842 | 2.015 | 2.074 | 2.120 | 2.098 |
| 0.52 | 2.079 | 2.192 | 2.254 | 2.310 | 2.243 |
| 0.53 | 1.937 | 2.188 | 2.211 | 2.301 | 2.468 |
| 0.54 | 1.494 | 1.627 | 1.688 | 1.744 | 1.813 |
| 0.55 | 1.561 | 1.689 | 1.682 | 1.693 | 1.533 |
| 0.56 | 1.786 | 1.808 | 1.780 | 1.777 | 1.695 |
| 0.57 | 0.938 | 0.969 | 0.959 | 0.960 | 0.966 |
| 0.58 | 1.184 | 1.191 | 1.190 | 1.190 | 1.206 |
| 0.59 | 1.015 | 1.006 | 1.008 | 1.009 | 1.027 |

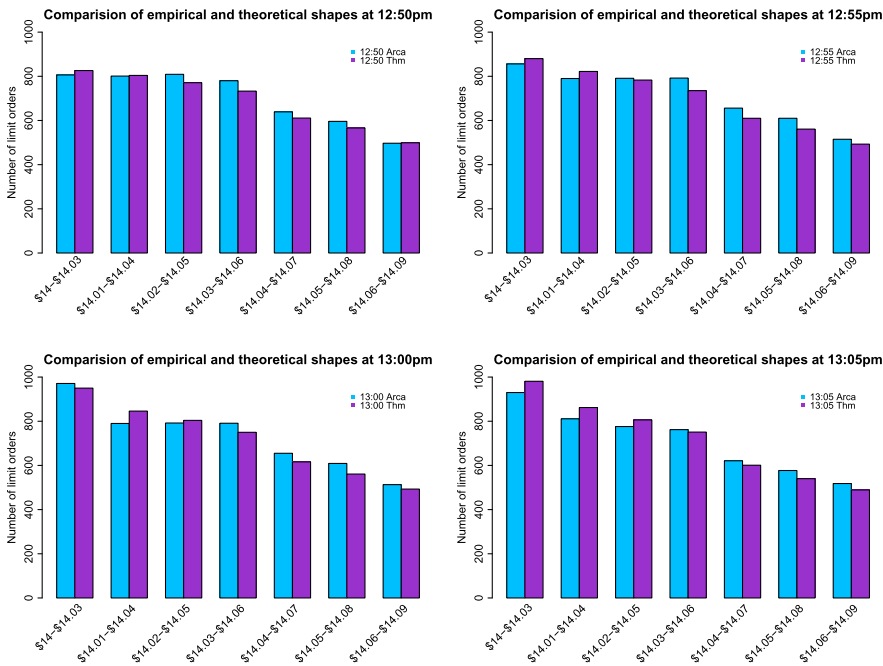


FIGURE 1. Comparisons of empirical sell-side shapes and theoretical sell-side shapes for BAC on August 5, 2010. Each blue bar corresponds to the total number of limit sell orders in each 3-cent price bin and it is obtained from data. Each magenta bar corresponds to the total number of limit sell orders in each price bin and it is obtained from our model. We have seven “consecutive” 3-cent price bins in each subfigure since we focus on price levels 10 ticks above the best ask price. The unit order size is taken to be 363-shares, which is the average limit sell order size of BAC during 12:45–13:05 pm. The best ask prices are the same (\$14) at four time instants 12:50, 12:55, 13:00, and 13:05 pm. Each time instant (say, 12:50 pm) represents the time that the first order book event occurs after that time instant.

For convenience, the aggregated order volume in each price bin is scaled by 363-shares, the average limit sell order size of BAC during 12:45–13:05 pm. From Figure 1, we observe good agreement between the theoretical shapes computed from our model (Theorem 3.1) and the empirical shapes of stock BAC on Arca. Thus, despite the simplifying assumptions, there is positive evidence that our limiting model can potentially explain the evolution of the shape of an order book over a time horizon that is large compared with the length of time between order book events.

Finally, we remark that the main goal of the empirical analysis is to illustrate the potential relevance of the fluid approximation rather than to make prediction about the future order-book shape, so no out-of-sample prediction result is reported here. In fact, we have attempted the analysis and found that the model does not work well for prediction purpose. One possible reason is that the parameters (order flow intensities) in our fluid model are constants, while in reality they vary over time and are random on the time scale of several minutes. This begs for further research on predictive models of order flow intensities and order-book shape, which is outside the scope of this paper.

5. PRELIMINARIES

This section presents preliminaries for proving the main result (Theorem 3.1).

5.1. A Technical Lemma

In this section, we introduce a technical lemma. It connects the process $(\zeta^{n,+}, \zeta^{n,-})$ with the process ζ^n which plays a fundamental role in the proof of Theorem 3.1.

Note that for fixed $f \in C[0, 1]$, one can construct a family of functions $f_\gamma \in C[0, 1]$ indexed by $\gamma \in (0, p)$ such that $\{f_\gamma\}$ are uniformly bounded and

$$f_\gamma(x) = \begin{cases} f(x) & \text{if } p \leq x \leq 1, \\ 0 & \text{if } 0 \leq x \leq p - \gamma, \\ \text{smooth} & \text{if } p - \gamma < x < p. \end{cases} \tag{5.1}$$

It is clear that

$$\lim_{\gamma \rightarrow 0^+} f_\gamma(x) = f(x) \cdot 1_{[p,1]}(x) \quad \text{for each } x \in [0, 1]. \tag{5.2}$$

The following lemma says that $\langle \zeta_t^{n,+}, f \rangle$ is “close” to $\langle \zeta_t^n, f_\gamma \rangle$ and $\langle \zeta_t^{n,-}, f \rangle$ is “close” to $\langle \zeta_t^n, f_\gamma - f \rangle$ when n is large and γ is small. The proof is given in Appendix A.

LEMMA 5.1: *Fix a function $f \in C[0, 1]$. There exists a positive constant C depending on T and two non-negative sequences $\{\tilde{a}_n\}, \{\hat{a}_n\}$ both depending on T and both approaching 0 as $n \rightarrow \infty$, such that for each $\gamma \in (0, p)$ and n large, we have*

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |\langle \zeta_t^{n,+}, f \rangle - \langle \zeta_t^n, f_\gamma \rangle| \right] \leq \tilde{a}_n + C\gamma, \tag{5.3}$$

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |\langle \zeta_t^{n,-}, f \rangle - \langle \zeta_t^n, f_\gamma - f \rangle| \right] \leq \hat{a}_n + C\gamma. \tag{5.4}$$

5.2. Tightness of $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$

In this section, we define the Polish space $\mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1]))$ on which the pair of measure-valued processes $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$ lives and discuss the tightness of this sequence.

Recall that the space of finite *non-negative* measures on $[0, 1]$, denoted by $\mathcal{M}^+([0, 1])$, is a Polish space under the following metric d_+ :

$$d_+(v_\alpha, v_\beta) = \sum_{k=1}^\infty \frac{1}{2^k} \frac{|\langle v_\alpha, \phi_k \rangle - \langle v_\beta, \phi_k \rangle|}{1 + |\langle v_\alpha, \phi_k \rangle - \langle v_\beta, \phi_k \rangle|}, \tag{5.5}$$

where $v_\alpha, v_\beta \in \mathcal{M}^+([0, 1])$, $\{\phi_k : k \geq 1\}$ are chosen to be a dense subset of $C([0, 1])$ endowed with uniform topology, and $\langle v, \phi_k \rangle \equiv \int \phi_k(x)v(dx)$ for measure v , see, for example, [18, Section 4.1]. Relying on the metric d_+ , we can define a metric d on the product space $\mathcal{M}^+([0, 1]) \times \mathcal{M}^+([0, 1])$ such that

$$d\left((v_\alpha^+, v_\alpha^-), (v_\beta^+, v_\beta^-)\right) \triangleq \sqrt{d_+^2(v_\alpha^+, v_\beta^+) + d_+^2(v_\alpha^-, v_\beta^-)}, \tag{5.6}$$

where $(v_\alpha^+, v_\alpha^-), (v_\beta^+, v_\beta^-) \in \mathcal{M}^+([0, 1]) \times \mathcal{M}^+([0, 1])$. The product space $\mathcal{M}^+([0, 1]) \times \mathcal{M}^+([0, 1])$ equipped with metric d defined in (5.6) is denoted by $\overline{\mathcal{M}}([0, 1])$. One readily verifies that $\overline{\mathcal{M}}([0, 1])$ is a Polish space (see, e.g., [20]). Therefore, the Skorohod space $\mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1]))$ on which $(\zeta^{n,+}, \zeta^{n,-})$ lives is also a Polish space with Skorohod J_1 topology. See [4] for more details on the topology.

We now state the main result of this section. The proof is given in Appendix B. Recall a sequence of random elements is called *C-tight* if it is tight, and if all the limit points of the sequence are concentrated on continuous paths. See, for example, [17, Definition VI.3.25].

PROPOSITION 5.1: *Fix a function $f \in C[0, 1]$. Then $\{(\langle \zeta^{n,+}, f \rangle, \langle \zeta^{n,-}, f \rangle) : n \geq 1\}$ is C-tight in $\mathbb{D}([0, T], \mathbb{R}^2)$, and $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$ is C-tight in $\mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1]))$.*

Proposition 5.1 together with Prohorov’s Theorem implies that the sequence $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$ is relatively compact. The next step is to characterize the limit points of this relatively compact sequence. Suppose (ζ^+, ζ^-) is a limit point, that is, there is a subsequence $\{(\zeta^{n_k,+}, \zeta^{n_k,-}) : k = 1, 2 \dots\}$ such that

$$(\zeta^{n_k,+}, \zeta^{n_k,-}) \Rightarrow (\zeta^+, \zeta^-) \text{ in } \mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1])) \text{ as } n_k \rightarrow \infty. \tag{5.7}$$

To characterize the pair (ζ^+, ζ^-) , we rely on an auxiliary process defined by

$$\zeta_t \triangleq \zeta_t^+ - \zeta_t^- \text{ for each } t \in [0, T], \tag{5.8}$$

5.3. Characterization of the Process ζ

In this section, we characterize the process ζ defined in (5.8).

It is straightforward to verify that for $f \in C[0, 1]$, the function from $(\pi^+, \pi^-) \in \mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1]))$ to $(\langle \pi^+, f \rangle, \langle \pi^-, f \rangle) \in \mathbb{D}([0, T], \mathbb{R}^2)$ is continuous with respect to Skorohod J_1 topology [use (5.5) and (5.6) and Lemma A.24 in [30]]. Now (5.7) together with continuous mapping theorem implies that for an arbitrary fixed function $f \in C([0, 1])$ we

have as $n_k \rightarrow \infty$,

$$(\langle \zeta^{n_k,+}, f \rangle, \langle \zeta^{n_k,-}, f \rangle) \Rightarrow (\langle \zeta^+, f \rangle, \langle \zeta^-, f \rangle) \quad \text{in } \mathbb{D}([0, T], \mathbb{R}^2). \tag{5.9}$$

Thus, we deduce that

$$\langle \zeta^{n_k}, f \rangle \Rightarrow \langle \zeta, f \rangle \quad \text{in } \mathbb{D}([0, T], \mathbb{R}), \tag{5.10}$$

where ζ^n is given in (3.3) and ζ is given in (5.8).

Next, for fixed $n \geq 1$, if we define a function F on \mathbb{R}^n by

$$F(x_1, \dots, x_n) = \frac{1}{n^2} \sum_{k=1}^n f\left(\frac{k}{n}\right) x_k, \tag{5.11}$$

we obtain from (3.3) and Markov property of \mathcal{X}^n that for $t \in [0, T]$,

$$\begin{aligned} \langle \zeta_t^n, f \rangle &= \frac{1}{n^2} \sum_{i=1}^n \mathcal{X}_i^n(nt) \cdot f\left(\frac{i}{n}\right) = F(\mathcal{X}^n(nt)) \\ &= \langle \zeta_0^n, f \rangle + \int_0^{nt} \mathcal{L}_n F(\mathcal{X}^n(s)) ds + M_t^n. \end{aligned} \tag{5.12}$$

Here \mathcal{L}_n is operator given in (2.5) and M^n is a (local) martingale. The martingale representation (5.12) and the convergence in (5.10) are the keys to characterize ζ .

To state the main result of this section, we recall the functions $\Lambda_A, \Lambda_B, \Theta_A$, and Θ_B given in Assumption 2.2 and \mathfrak{p} given in Assumption 2.1. For notational convenience, we let ν_Λ be a signed measure that is absolutely continuous with respect to Lebesgue measure such that for $x \in [0, 1]$

$$\frac{\nu(dx)}{dx} = \Lambda_A(x - \mathfrak{p}) \cdot 1_{x > \mathfrak{p}} - \Lambda_B(\mathfrak{p} - x) \cdot 1_{x < \mathfrak{p}}. \tag{5.13}$$

In addition, we let \mathcal{A}_Θ be a linear operator such that for $f \in C([0, 1])$,

$$\mathcal{A}_\Theta f(x) = f(x) \cdot \left(\Theta_A(x - \mathfrak{p}) \cdot 1_{x > \mathfrak{p}} + \Theta_B(\mathfrak{p} - x) \cdot 1_{x < \mathfrak{p}} \right). \tag{5.14}$$

We now state the main result of this section which gives a characterization of ζ .

PROPOSITION 5.2: *Let (ζ^+, ζ^-) be a limit point in (5.7) and let ζ be defined in (5.8). Then ζ satisfies the following set of equations: for any $f \in C([0, 1])$ and $t \in [0, T]$,*

$$\langle \zeta_t, f \rangle = \langle \zeta_0, f \rangle + \langle \nu, f \rangle \cdot t - \int_0^t \langle \zeta_s, \mathcal{A}_\Theta f \rangle ds, \tag{5.15}$$

where ν and \mathcal{A}_Θ are given in (5.13) and (5.14). Moreover, $\langle \zeta, f \rangle$ has continuous trajectory and there is a unique deterministic measure-valued process solving Eq. (5.15). Finally, for fixed $t \in [0, T]$, ζ_t is a finite measure that is absolutely continuous with respect to Lebesgue measure, and it has bounded density function φ determined by (3.6)–(3.9).

The main technical part of the proof of Proposition 5.2 consists of showing that in (5.12) as $n \rightarrow \infty$, the martingale term vanishes and the term involving the generator converges weakly. We leave the details of the proof to Appendix C.

6. PROOF OF THEOREM 3.1

In this section, we prove the main result Theorem 3.1. The proof of part (b) of Theorem 3.1 relies on the results in Section 5 while the proof of part (a) does not.

6.1. Proof of Part (a) of Theorem 3.1

PROOF: We prove (3.4) using a stochastic comparison method. The convergence in (3.5) follows from a similar argument and is thus omitted.

We first note that

$$\lim_{n \rightarrow \infty} \frac{p_A^n(0)}{n} = \lim_{n \rightarrow \infty} \frac{p_B^n(0)}{n} = p. \tag{6.1}$$

This readily follows from the definition of best bid and best ask prices and the regularity condition (b) of ϱ in Assumption 2.1. Hence, in order to prove (3.4), it suffices to show

$$\sup_{0 \leq t \leq T} \left| \frac{p_A^n(nt)}{n} - \frac{p_A^n(0)}{n} \right| \Rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Given any $\epsilon > 0$, we have

$$\begin{aligned} & \mathbb{P} \left(\sup_{0 \leq t \leq T} \left| \frac{p_A^n(nt)}{n} - \frac{p_A^n(0)}{n} \right| > \epsilon \right) \\ & \leq \mathbb{P} \left(\sup_{0 \leq t \leq T} p_A^n(nt) - p_A^n(0) > n\epsilon \right) + \mathbb{P} \left(p_A^n(0) - \inf_{0 \leq t \leq T} p_A^n(nt) > n\epsilon \right). \end{aligned} \tag{6.2}$$

We next show that for any small $\delta > 0$, there exists N_δ such that when $n > N_\delta$,

$$\mathbb{P} \left(\sup_{0 \leq t \leq T} p_A^n(nt) - p_A^n(0) > n\epsilon \right) \leq \delta. \tag{6.3}$$

Since $p_A^n(\cdot)$ is upper bounded by $n + 1$ by definition, we deduce that if $p_A^n(0) \geq n + 1 - n\epsilon$, then (6.3) follows trivially. Thus below we assume $p_A^n(0) < n + 1 - n\epsilon$, or equivalently, $p_A^n(0) + \lfloor n\epsilon \rfloor \leq n$. For a fixed n and a given $p_A^n(0)$, we first define a map g from \mathbb{Z}^n to \mathbb{Z} :

$$g(x) = \sum_{i=p_A^n(0)}^{p_A^n(0)+\lfloor n\epsilon \rfloor} x_i^+. \tag{6.4}$$

We use the map g to further introduce an auxiliary process \mathcal{Z}^n to track the number of sell limit orders on price levels from $p_A^n(0)$ to $p_A^n(0) + \lfloor n\epsilon \rfloor$. That is, for each $t \geq 0$,

$$\mathcal{Z}^n(t) \triangleq g(\mathcal{X}^n(nt)) = \sum_{i=p_A^n(0)}^{p_A^n(0)+\lfloor n\epsilon \rfloor} [\mathcal{X}_i^n(nt)]^+. \tag{6.5}$$

Write

$$\sigma_{\mathcal{Z}^n}^n \triangleq \inf \{ t \geq 0 : \mathcal{Z}^n(t) = 0 \}. \tag{6.6}$$

Then $\sigma_{\mathcal{Z}^n}^n$ is the first time that \mathcal{Z}^n reaches 0 starting from $\mathcal{Z}^n(0) > 0$, where $\mathcal{Z}^n(0) > 0$ follows from the fact that $p_A^n(0) \leq n$. Thus, in order to show (6.3), it suffices to prove

$$\mathbb{P}(\sigma_{\mathcal{Z}^n}^n \leq T) \leq \delta, \quad \text{when } n > N_\delta. \tag{6.7}$$

The key idea for proving (6.7) is to construct a n -dimensional CTMC $\{\mathcal{W}^n(t) : t \geq 0\}$ on \mathbb{Z}^n that is dominated by $\{\mathcal{X}^n(t) : t \geq 0\}$ in the sense of strong stochastic order for each

n . That is, $\mathbb{E}r(\mathcal{W}^n(t)) \leq \mathbb{E}r(\mathcal{X}^n(t))$ for all real-valued increasing functions r on \mathbb{Z}^n and all $t \geq 0$. To proceed, we set $\mathcal{W}^n(0) = \mathcal{X}^n(0)$. For convenience, we define two constants:

$$\bar{\Lambda} \triangleq 2 \cdot \left[\left(\max_{x \in [0,1]} \Lambda_A(x) \right) \vee \left(\max_{x \in [0,1]} \Lambda_B(x) \right) \right] < \infty, \tag{6.8}$$

$$\bar{\Theta} \triangleq \left(\max_{x \in [0,1]} \Theta_A(x) \right) \vee \left(\max_{x \in [0,1]} \Theta_B(x) \right) > 0. \tag{6.9}$$

We construct \mathcal{W}^n by modifying the transition parameters of the CTMC \mathcal{X}^n as follows: (a) set the limit sell order arrival rate to zero; (b) set the limit buy order cancellation rate and the market sell order arrival rate to zero; (c) set the limit buy order arrival rate at each price level to $\bar{\Lambda}$ given in (6.8), the market buy order arrival rate to $n^\kappa \cdot \Upsilon$ given in Assumption 2.2, and the limit sell order cancellation rate per order to $\frac{1}{n} \bar{\Theta}$.

We next argue \mathcal{W}^n is stochastically dominated by \mathcal{X}^n . We apply the result of [27]. Recall a set $\Gamma \subset \mathbb{Z}^n$ is said to be decreasing if $x \in \Gamma$ implies $\{y \in \mathbb{Z}^n : y \leq x\} \subset \Gamma$. From the construction of \mathcal{W}^n , we know that there are no limit sell order arrivals and limit buy order removals (due to cancellation or matches with market sell orders) in \mathcal{W}^n . This implies that if \mathcal{W}^n transits from one state $w \in \mathbb{Z}^n$ to another state $w' \in \mathbb{Z}^n$, then $w' \leq w$ with respect to the partial order \leq on \mathbb{Z}^n . Thus we deduce from [27, Theorem 5.3] that it suffices to show for all decreasing sets $\Gamma \subset \mathbb{Z}^n$ such that $w \notin \Gamma$

$$\sum_{z \in \Gamma} q_{\mathcal{W}^n}(w, z) \geq \sum_{z \in \Gamma} q_{\mathcal{X}^n}(w, z) \quad \text{for all } w \leq x, \tag{6.10}$$

where $q_{\mathcal{W}^n}$ and $q_{\mathcal{X}^n}$ are transition rates for \mathcal{W}^n and \mathcal{X}^n . One readily verifies (6.10) from the construction of \mathcal{W}^n and the fact that both \mathcal{W}^n and \mathcal{X}^n have the property that only one component can change its value at each transition.

Given that \mathcal{W}^n is stochastically dominated by \mathcal{X}^n , we are ready to construct a process $\tilde{\mathcal{Z}}^n$ that is stochastically dominated by \mathcal{Z}^n given in (6.5). We simply set for $t \geq 0$

$$\tilde{\mathcal{Z}}^n(t) \triangleq g(\mathcal{W}^n(nt)) = \sum_{i=p_A^n(0)}^{p_A^n(0)+[n\epsilon]} [\mathcal{W}_i^n(nt)]^+ \geq 0. \tag{6.11}$$

Then \mathcal{Z}^n dominates $\tilde{\mathcal{Z}}^n$ in strong stochastic order since g in (6.4) is an increasing function on \mathbb{Z}^n . We then deduce that

$$\mathbb{P}(\sigma_{\mathcal{Z}^n}^n \leq T) \leq \mathbb{P}(\sigma_{\tilde{\mathcal{Z}}^n}^n \leq T),$$

where $\sigma_{\tilde{\mathcal{Z}}^n}^n$ is the first time that $\tilde{\mathcal{Z}}^n$ reaches zero starting from $\tilde{\mathcal{Z}}^n(0) = \mathcal{Z}^n(0)$, that is,

$$\sigma_{\tilde{\mathcal{Z}}^n}^n \triangleq \inf\{t \geq 0 : \tilde{\mathcal{Z}}^n(t) = 0\}. \tag{6.12}$$

Hence, to show (6.7) it suffices to prove that given any $\delta > 0$, there exists N_δ such that

$$\mathbb{P}(\sigma_{\mathcal{Z}^n}^n \leq T) \leq \delta, \quad \text{when } n > N_\delta. \tag{6.13}$$

We now focus on proving (6.13). One can verify that the process $\tilde{\mathcal{Z}}^n$ defined in (6.11) is a pure-death process with absorbing barrier at zero, and the death rate of $\tilde{\mathcal{Z}}^n$ when it is

in state $k \in \{1, 2, \dots\}$ is given by

$$n \cdot \left(k \cdot \frac{\bar{\Theta}}{n} + n^\kappa \cdot \Upsilon \right). \tag{6.14}$$

One can also check from Assumption 2.1, (6.11) and (6.1) that

$$\lim_{n \rightarrow \infty} \frac{\tilde{Z}^n(0)}{n^2} = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=p_A^n(0)}^{p_A^n(0)+\lfloor n\epsilon \rfloor} [\mathcal{X}_i^n(0)]^+ = \int_p^{p+\epsilon} \varrho(x) dx > 0. \tag{6.15}$$

Since \tilde{Z}^n is a pure death process with absorbing state zero, we obtain

$$\sigma_{\tilde{Z}^n}^n = \sum_{k=1}^{\tilde{Z}^n(0)} D_k, \tag{6.16}$$

where D_k represents the first passage time that \tilde{Z}^n starts from state k and reaches $k - 1$, so D_k is an exponential random variable with rate $k \cdot \bar{\Theta} + n^{1+\kappa} \Upsilon$ given in (6.14). In addition, all the D_k 's are mutually independent. Thus, we have

$$\begin{aligned} \mathbb{E}[\sigma_{\tilde{Z}^n}^n] &= \mathbb{E} \left[\sum_{k=1}^{\tilde{Z}^n(0)} D_k \right] = \sum_{k=1}^{\tilde{Z}^n(0)} \frac{1}{k \cdot \bar{\Theta} + n^{1+\kappa} \cdot \Upsilon}, \\ \text{Var}(\sigma_{\tilde{Z}^n}^n) &= \sum_{k=1}^{\tilde{Z}^n(0)} \text{Var}(D_k) = \sum_{k=1}^{\tilde{Z}^n(0)} \frac{1}{(k \cdot \bar{\Theta} + n^{1+\kappa} \cdot \Upsilon)^2}. \end{aligned}$$

Since $\kappa < 1$, we deduce from (6.15) that there exists C independent of n such that

$$\lim_{n \rightarrow \infty} \frac{1}{\ln n} \mathbb{E}[\sigma_{\tilde{Z}^n}^n] = \frac{1 - \kappa}{\bar{\Theta}} > 0, \quad \text{and} \quad \sup_n \text{Var}(\sigma_{\tilde{Z}^n}^n) \leq C. \tag{6.17}$$

Thus, for all n large enough, we obtain $\mathbb{E}[\sigma_{\tilde{Z}^n}^n] > T$. It follows that

$$\begin{aligned} \text{Var}(\sigma_{\tilde{Z}^n}^n) &\geq \mathbb{E} \left[(\mathbb{E}[\sigma_{\tilde{Z}^n}^n] - \sigma_{\tilde{Z}^n}^n)^2 : \sigma_{\tilde{Z}^n}^n \leq T \right] \\ &\geq \mathbb{E} \left[(\mathbb{E}[\sigma_{\tilde{Z}^n}^n] - T)^2 : \sigma_{\tilde{Z}^n}^n \leq T \right] \\ &= (\mathbb{E}[\sigma_{\tilde{Z}^n}^n] - T)^2 \cdot \mathbb{P}(\sigma_{\tilde{Z}^n}^n \leq T). \end{aligned}$$

Therefore we deduce that when n is large,

$$\mathbb{P}(\sigma_{\tilde{Z}^n}^n \leq T) \leq \frac{1}{(\mathbb{E}[\sigma_{\tilde{Z}^n}^n] - T)^2} \text{Var}(\sigma_{\tilde{Z}^n}^n).$$

This yields (6.13) after applying (6.17). Thus we have completed the proof of (6.3).

We next show that for n large,

$$\mathbb{P}\left(p_A^n(0) - \inf_{0 \leq t \leq T} p_A^n(nt) > n\epsilon\right) \leq \delta. \tag{6.18}$$

Note that $p_B^n(\cdot)$ is smaller than $p_A^n(\cdot)$ at each time, we have

$$\begin{aligned} p_A^n(0) - \inf_{0 \leq t \leq T} p_A^n(nt) &\leq p_A^n(0) - \inf_{0 \leq t \leq T} p_B^n(nt) \\ &= p_B^n(0) - \inf_{0 \leq t \leq T} p_B^n(nt) + p_A^n(0) - p_B^n(0). \end{aligned}$$

Now (6.1) implies that for n large,

$$\mathbb{P}\left(p_A^n(0) - p_B^n(0) > \frac{n\epsilon}{2}\right) \leq \frac{\delta}{2}.$$

Thus, it suffices to prove that for n large

$$\mathbb{P}\left(p_B^n(0) - \inf_{0 \leq t \leq T} p_B^n(nt) > \frac{n\epsilon}{2}\right) \leq \frac{\delta}{2}.$$

This follows from a similar argument for (6.3). Thus, we have (6.18). On combining (6.2), (6.3), and (6.18), we obtain (3.4). Therefore the proof is complete. ■

6.2. Proof of Part (b) of Theorem 3.1

In this section, we prove part (b) of Theorem 3.1. We rely on the results in Section 5 and characterize the limit point (ζ^+, ζ^-) in (5.7).

PROOF: Assume (ζ^+, ζ^-) is a limit point as in Proposition 5.2. In view of Proposition 5.2, to show that ζ_t^+ (respectively ζ_t^-) is absolutely continuous with density $\varphi^+(\cdot, t)$ (respectively $\varphi^-(\cdot, t)$), it suffices to show that for any $f \in C([0, 1])$ and $t \in [0, T]$, we have

$$\langle \zeta_t^+, f \rangle = \int_0^1 f(x)\varphi^+(x, t)dx, \tag{6.19}$$

$$\langle \zeta_t^-, f \rangle = \int_0^1 f(x)\varphi^-(x, t)dx, \tag{6.20}$$

where $\varphi^+ = \max\{\varphi, 0\}$, $\varphi^- = \max\{-\varphi, 0\}$ and φ is uniquely determined by (3.6)–(3.9).

To this end, we first note that $\langle \zeta_t^{n_k, +}, f \rangle \Rightarrow \langle \zeta_t^+, f \rangle$ as $n_k \rightarrow \infty$ for fixed time $t \in [0, T]$. This is true because $\langle \zeta_t^{n_k, +}, f \rangle \Rightarrow \langle \zeta^+, f \rangle$ in (5.9) and almost every path of the limit point $\langle \zeta^+, f \rangle$ is continuous by Proposition 5.1. Similarly, we deduce from (5.10) that $\langle \zeta_t^{n_k}, f_\gamma \rangle \Rightarrow \langle \zeta_t, f_\gamma \rangle$ for $f_\gamma \in C[0, 1]$ introduced in (5.1) and $t \in [0, T]$. Since ζ is deterministic by Proposition 5.2, we obtain that $\langle \zeta_t^{n_k}, f_\gamma \rangle$ converges in probability to $\langle \zeta_t, f_\gamma \rangle$. Therefore, we have for fixed $\gamma > 0$,

$$\langle \zeta_t^{n_k, +}, f \rangle - \langle \zeta_t^{n_k}, f_\gamma \rangle \Rightarrow \langle \zeta_t^+, f \rangle - \langle \zeta_t, f_\gamma \rangle \quad \text{as } n_k \rightarrow \infty.$$

Using Theorem 3.4 in [4], we obtain from Lemma 5.1 that there exists a constant $C > 0$ such that for each $\gamma \in (0, \mathfrak{p})$ and $t \in [0, T]$

$$\begin{aligned} \mathbb{E} \left| \langle \zeta_t^+, f \rangle - \langle \zeta_t, f_\gamma \rangle \right| &\leq \liminf_{n_k \rightarrow \infty} \mathbb{E} \left| \langle \zeta_t^{n_k, +}, f \rangle - \langle \zeta_t^{n_k}, f_\gamma \rangle \right| \\ &\leq \liminf_{n_k \rightarrow \infty} (\tilde{a}_{n_k} + C\gamma) \\ &= C\gamma. \end{aligned} \tag{6.21}$$

In addition, since ζ_t is a deterministic signed measure with bounded density function by Proposition 5.2 and $\{f_\gamma : \gamma \in (0, \mathfrak{p})\}$ is uniformly bounded on $[0, 1]$, one infers from the dominated convergence theorem and (5.2) that

$$\lim_{\gamma \rightarrow 0^+} \langle \zeta_t, f_\gamma \rangle = \langle \zeta_t, \lim_{\gamma \rightarrow 0^+} f_\gamma \rangle = \int_{\mathfrak{p}}^1 f(x)\varphi(x, t)dx. \tag{6.22}$$

Since φ satisfies (3.6)-(3.9), and ρ satisfies (2.7), we deduce that φ is non-negative on $[\mathfrak{p}, 1]$ and non-positive on $[0, \mathfrak{p}]$. This implies

$$\int_{\mathfrak{p}}^1 f(x)\varphi(x, t)dx = \int_0^1 f(x)\varphi^+(x, t)dx. \tag{6.23}$$

Therefore, we find from (6.21) that

$$\begin{aligned} \mathbb{E} \left| \langle \zeta_t^+, f \rangle - \int_0^1 f(x)\varphi^+(x, t)dx \right| &\leq \mathbb{E} \left| \langle \zeta_t^+, f \rangle - \langle \zeta_t, f_\gamma \rangle \right| + \left| \langle \zeta_t, f_\gamma \rangle - \int_0^1 f(x)\varphi^+(x, t)dx \right| \\ &\leq C\gamma + \left| \langle \zeta_t, f_\gamma \rangle - \int_0^1 f(x)\varphi^+(x, t)dx \right|. \end{aligned} \tag{6.24}$$

Let $\gamma \rightarrow 0^+$ in inequality (6.24), we conclude from (6.22) and (6.23) that

$$\mathbb{E} \left| \langle \zeta_t^+, f \rangle - \int_0^1 f(x)\varphi^+(x, t)dx \right| = 0.$$

Thus, (6.19) holds with probability one. A similar argument yields (6.20). The proof is therefore complete. ■

Acknowledgement

We are grateful to the editor and the anonymous referee for helpful suggestions, which improve the exposition of this paper. We thank Jim Dai, Ton Dieker, Rama Cont, and Yu Gu for helpful discussions. Xuefeng Gao acknowledges support from Hong Kong RGC ECS Grant 24207015, and CUHK Direct Grants 4055035 and 4055054. S.J. Deng is supported in part by a PSERC grant.

References

1. Abergel, F. & Jedidi, A. (2013). On the stability and price scaling limit of a Hawkes process-based order book model. Working paper. Available at <http://ssrn.com/abstract=2263162>.
2. Aldous, D. (1981). Weak convergence and the general theory of processes. Incomplete draft of a monograph, <http://www.stat.berkeley.edu/aldous/Papers/weak-gtp.pdf>.
3. Bayer, C., Horst, U., & Qiu, J. (2014). A functional limit theorem for limit order books with state dependent price dynamics. arXiv preprint arXiv:1405.5230.
4. Billingsley, P. (1999). *Convergence of probability measures*. 2nd ed. New York: Wiley.
5. Blanchet, J. & Chen, X. (2013). Continuous-time modeling of bid-ask spread and price dynamics in limit order books. arXiv preprint arXiv:1310.1103.

6. Cont, R. (2015). High frequency dynamics of limit order markets: Multi-scale modeling and asymptotic analysis. Working paper. Slides available at <https://workspace.imperial.ac.uk/quantitativefinance/Public/events/SLIDES%203RD%20IMPERIAL%20ETH/Rama%20Cont%20.pdf>
7. Cont, R. & De Larrard, A. (2011). Order book dynamics in liquid markets: Limit theorems and diffusion approximations. Working Paper. Available at <http://ssrn.com/abstract=1757861>
8. Cont, R. & De Larrard, A. (2013). Price dynamics in a Markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1): 1–25.
9. Cont, R., Stoikov, S., & Talreja, R. (2010). A stochastic model for order book dynamics. *Operations Research*, 58(3): 549–563.
10. Dawson, D. (1993). *Measure-valued Markov processes*. Berlin, Heidelberg: Springer.
11. Driver, K.B. (2003). Analysis tools with applications. *Lecture Notes*. Available at http://www.math.ucsd.edu/~bdriver/240-01-02/Lecture_Notes/anal2p.pdf
12. Gould, M.D., Porter, M.A., Williams, S., McDonald, M., Fenn, D.J., & Howison, S.D. (2013). Limit order books. *Quantitative Finance*, 13(11): 1709–1742.
13. Hautsch, N. & Huang, R. (2011). Limit order flow, market impact and optimal order sizes: Evidence from NASDAQ TotalView–ITCH data. Working Paper. Available at <http://ssrn.com/abstract=1914293>
14. Horst, U. & Kreher, D. (2015). A weak law of large numbers for a limit order book model with fully state dependent order dynamics. arXiv preprint arXiv:1502.04359.
15. Horst, U. & Paulsen, M. (2015). A law of large numbers for limit order books. arXiv preprint arXiv:1501.00843.
16. Huang, H. & Kercheval, A.N. (2012). A generalized birth–death stochastic model for high-frequency order book dynamics. *Quantitative Finance*, 12(4): 547–557.
17. Jacod, J. & Shiryaev, A.N. (2003). *Limit Theorems for stochastic processes*. 2nd ed. Berlin: Springer.
18. Kipnis, C. & Landim, C. (1999). *Scaling limits of interacting particle systems*, Vol. 320. Berlin, Heidelberg: Springer.
19. Kirilenko, A., Sowers, R.B., & Meng, X. (2013). A multiscale model of high-frequency trading. *Algorithmic Finance*, 2(1): 59–98.
20. Kotelenetz, P. (2010). Stochastic flows and signed measure valued stochastic partial differential equations. *Theory of Stochastic Process*, 16(2): 86–105.
21. Kruk, L. (2003). Functional limit theorems for a simple auction. *Mathematics of Operations Research*, 28(4): 716–751.
22. Kruk, L. (2012). Limiting distribution for a simple model of order book dynamics. *Central European Journal of Mathematics*, 10(6): 2283–2295.
23. Kurtz, T.G. (1971). Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability*, 8(2): 344–356.
24. Lakner, P., Reed, J., & Simatos, F. (2015). Scaling limits for a limit order book model using the regenerative characterization of Levy trees. <http://people.stern.nyu.edu/jreed/papers/paper20.pdf>
25. Lakner, P., Reed, J., & Stoikov, S. (2016). High frequency asymptotics for the limit order book. Market Microstructure and Liquidity. To appear. <http://people.stern.nyu.edu/jreed/papers/paper19.pdf>
26. Maglaras, C., Moallemi, C.C., & Zheng, H. (2015). Optimal execution in a limit order book and an associated microstructure market impact model. Available at <https://ssrn.com/abstract=2610808>
27. Massey, W.A. (1987). Stochastic orderings for Markov processes on partially ordered spaces. *Mathematics of Operations Research*, 12(2): 350–367.
28. Parlour, C.A. & Seppi, D.J. (2008). Limit order markets: A survey. *Handbook of financial intermediation and banking*, 5: 63–95.
29. Protter, P.E. (2004). *Stochastic integration and differential equations: Version 2.1*, Vol. 21. Berlin, Heidelberg: Springer.
30. Seppäläinen, T. (2008). Translation invariant exclusion processes. *Lecture notes*. Available at <http://www.math.wisc.edu/~seppalai/excl-book/ajo.pdf>.
31. Spiliopoulos, K., Sirignano, J.A., & Giesecke, K. (2014). Fluctuation analysis for the loss from default. *Stochastic Processes and their Applications*, 124(7): 2322–2362.
32. Tao, T. (2006). *Nonlinear dispersive and wave equations: Local and global analysis*. Number 106. American Mathematical Society.
33. Whitt, W. (2002). *Stochastic-process limits*. New York: Springer.

APPENDIX A. PROOF OF LEMMA 5.1

In this section, we prove Lemma 5.1 in Section 5.1. To do so, we first introduce a result on bounding the (scaled) total number of limit orders within a certain price range of the order book. The proof is deferred to the end of this appendix.

LEMMA A.1: Fix $T > 0$. There exists a positive constant C , which depends on T , but is independent of n such that for any fixed interval $[a, b] \subset [0, 1]$, we have

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq t \leq T} \frac{1}{n^2} \sum_{i: \frac{i}{n} \in [a,b]} |\mathcal{X}_i^n(nt)| \right]^2 \leq C^2(b - a)^2, \tag{A.1}$$

and

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq t \leq T} \frac{1}{n^2} \sum_{i: \frac{i}{n} \in [a,b]} |\mathcal{X}_i^n(nt)| \right] \leq C(b - a). \tag{A.2}$$

In addition for any $\sigma \in [0, T]$ we have

$$\limsup_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq s \leq T - \sigma} \int_s^{s+\sigma} \frac{1}{n^2} \sum_{k=1}^n |\mathcal{X}_k^n(nu)| du \right] \leq C\sigma. \tag{A.3}$$

We now prove Lemma 5.1.

PROOF OF LEMMA 5.1: We focus on proving (5.3). The proof for (5.4) follows similarly and hence is omitted. Throughout the proof, we use a generic constant C that may depend on T and may vary from line to line, but C is independent of n .

From the definitions of $\zeta^{n,+}$, ζ^n and f_γ (see (3.1), (3.3) and (5.1)) we derive

$$\begin{aligned} & \mathbb{E} \left[\sup_{t \in [0, T]} \left| \langle \zeta_t^{n,+}, f \rangle - \langle \zeta_t^n, f_\gamma \rangle \right| \right] \\ &= \frac{1}{n^2} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i > p_B^n(nt)} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) - \sum_{i=1}^n \mathcal{X}_i^n(nt) f_\gamma\left(\frac{i}{n}\right) \right| \right] \\ &= \frac{1}{n^2} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i > p_B^n(nt)} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) - \sum_{i \geq n\mathfrak{p}} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) - \sum_{i \in (n\mathfrak{p} - n\gamma, n\mathfrak{p})} \mathcal{X}_i^n(nt) f_\gamma\left(\frac{i}{n}\right) \right| \right] \\ &\leq \frac{1}{n^2} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i > p_B^n(nt)} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) - \sum_{i \geq n\mathfrak{p}} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) \right| \right] \\ &\quad + \frac{1}{n^2} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i \in (n\mathfrak{p} - n\gamma, n\mathfrak{p})} \mathcal{X}_i^n(nt) f_\gamma\left(\frac{i}{n}\right) \right| \right]. \tag{A.4} \end{aligned}$$

We now bound the two terms after the inequality sign in (A.4). We first bound the second term, that is, the last display in (A.4). Since the family of functions f_γ are constructed such that f_γ are uniformly bounded by some constant C on $[0, 1]$ for all $\gamma \in (0, \mathfrak{p})$, the second term in (A.4) is bounded above by

$$\frac{C}{n^2} \cdot \mathbb{E} \left[\sup_{t \in [0, T]} \sum_{i \in (n\mathfrak{p} - n\gamma, n\mathfrak{p})} |\mathcal{X}_i^n(nt)| \right],$$

which is further bounded above by $C\gamma$ for large n by (A.2).

We next bound the first term, which we denote by \tilde{a}_n . It is clear that $\tilde{a}_n \geq 0$, so we only have to show that $\lim_{n \rightarrow \infty} \tilde{a}_n = 0$. We split it into two parts and study them separately. Fix some $\delta > 0$. Since f is continuous, it is bounded by some constant C on $[0, 1]$. We then deduce that

$$\begin{aligned} & \frac{1}{n^2} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i > p_B^n(nt)} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) - \sum_{i \geq np} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) \right| ; \sup_{t \in [0, T]} \left| \frac{p_B^n(nt)}{n} - p \right| \geq \delta \right] \\ & \leq 2C \cdot \mathbb{E} \left[\sup_{0 \leq s \leq T} \sum_{i=1}^n \frac{1}{n^2} |\mathcal{X}_i^n(ns)| ; \sup_{t \in [0, T]} \left| \frac{p_B^n(nt)}{n} - p \right| \geq \delta \right] \triangleq b_{n, \delta}. \end{aligned} \tag{A.5}$$

Now (A.1) in Lemma A.1 implies that the sequence $\{\sup_{0 \leq s \leq T} \sum_{k=1}^n \frac{1}{n^2} |\mathcal{X}_k^n(ns)| : n \geq 1\}$ is uniformly bounded in L^2 , thus $\{\sup_{0 \leq s \leq T} \sum_{k=1}^n \frac{1}{n^2} |\mathcal{X}_k^n(ns)| : n \geq 1\}$ is uniformly integrable. Thus, we obtain from part (a) of Theorem 3.1 and (A.5) that for any $\delta > 0$,

$$\lim_{n \rightarrow \infty} b_{n, \delta} = 0. \tag{A.6}$$

We next proceed to bound

$$\frac{1}{n^2} \mathbb{E} \left[\sup_{t \in [0, T]} \left| \sum_{i > p_B^n(nt)} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) - \sum_{i \geq np} \mathcal{X}_i^n(nt) f\left(\frac{i}{n}\right) \right| ; \sup_{t \in [0, T]} \left| \frac{p_B^n(nt)}{n} - p \right| < \delta \right].$$

We can choose small $\delta > 0$ such that $(p - \delta, p + \delta) \subset (0, 1)$. The above term is then bounded above by

$$\frac{C}{n^2} \cdot \mathbb{E} \left[\sup_{t \in [0, T]} \sum_{i \in (np - n\delta, np + n\delta)} |\mathcal{X}_i^n(nt)| \right],$$

which is bounded above by $C\delta$ for n large using (A.2). Upon combining the two parts we find for large n

$$0 \leq \tilde{a}_n \leq b_{n, \delta} + C\delta \quad \text{for sufficiently small } \delta > 0.$$

Fix $\delta > 0$ and let $n \rightarrow \infty$. We infer from (A.6) that

$$0 \leq \limsup_{n \rightarrow \infty} \tilde{a}_n \leq C\delta.$$

Let $\delta \rightarrow 0+$, we obtain $\lim_{n \rightarrow \infty} \tilde{a}_n = 0$. Therefore the proof is complete. ■

PROOF OF LEMMA A.1: It suffices to prove (A.1) since it is clear that (A.1) implies (A.2) by the Jensen’s inequality. Inequality (A.2) further leads to (A.3) by observing that

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq s \leq T - \sigma} \int_s^{s + \sigma} \frac{1}{n^2} \sum_{k=1}^n |\mathcal{X}_k^n(nu)| du \right] & \leq \mathbb{E} \left[\sup_{0 \leq s \leq T - \sigma} \int_s^{s + \sigma} \sup_{u \in [s, s + \sigma]} \frac{1}{n^2} \sum_{k=1}^n |\mathcal{X}_k^n(nu)| du \right] \\ & \leq \sigma \cdot \mathbb{E} \left[\sup_{0 \leq s \leq T - \sigma} \sup_{u \in [s, s + \sigma]} \frac{1}{n^2} \sum_{k=1}^n |\mathcal{X}_k^n(nu)| \right] \\ & = \sigma \cdot \mathbb{E} \left[\sup_{u \leq T} \frac{1}{n^2} \sum_{k=1}^n |\mathcal{X}_k^n(nu)| \right]. \end{aligned}$$

Thus, the rest of the proof focus on establishing (A.1). For fixed n and $i \in \{1, \dots, n\}$, we write $A_i^n(t)$ for the total number of new limit orders submitted at price level i/n by time t . One can

construct independent Poisson processes $\bar{A}_1^n, \dots, \bar{A}_n^n$ with rate $\bar{\Lambda}$ given in (6.8) such that for each i , \bar{A}_i^n is defined on the probability space which A_i^n lives in, and

$$\bar{A}_i^n(t, \omega) \geq A_i^n(t, \omega) \quad \text{for every sample path } \omega \text{ and every } t \geq 0. \tag{A.7}$$

Set

$$\bar{\mathcal{E}}^n(t) \triangleq \sum_{i: \frac{i}{n} \in [a, b]} \bar{A}_i^n(nt) \quad \text{and} \quad \mathcal{N}^n(t) = \sum_{i: \frac{i}{n} \in [a, b]} |\mathcal{X}_i^n(nt)|.$$

We immediately get that $\bar{\mathcal{E}}^n(\cdot)$ is a Poisson process with rate r_n where

$$r_n \triangleq n\bar{\Lambda} \cdot (\lfloor nb \rfloor - \lceil na \rceil + 1) \leq n\bar{\Lambda} \cdot (nb - na + 1), \tag{A.8}$$

and

$$\sup_{0 \leq t \leq T} \mathcal{N}^n(t) \leq \mathcal{N}^n(0) + \bar{\mathcal{E}}^n(T) \quad \text{for every sample path.}$$

This implies that

$$\mathbb{E}[\sup_{0 \leq t \leq T} \mathcal{N}^n(t)]^2 \leq \mathbb{E}[\mathcal{N}^n(0) + \bar{\mathcal{E}}^n(T)]^2 \leq 2\mathbb{E}[\mathcal{N}^n(0)]^2 + 2\mathbb{E}[\bar{\mathcal{E}}^n(T)]^2. \tag{A.9}$$

One readily checks from Assumption 2.1 that

$$\frac{1}{n^4} \mathbb{E}[\mathcal{N}^n(0)]^2 = \left(\frac{1}{n} \sum_{i: \frac{i}{n} \in [a, b]} |\varrho(i/n)| \right)^2 \leq \left[C \left((b-a) + \frac{1}{n} \right) \right]^2, \tag{A.10}$$

where we use the fact that ϱ is bounded by some constant C on $[0, 1]$. To bound $\mathbb{E}[\bar{\mathcal{E}}^n(T)]^2$ in (A.9), we note that $\bar{\mathcal{E}}^n$ is a Poisson process with rate r_n . Thus, we have

$$\mathbb{E}[\bar{\mathcal{E}}^n(T)]^2 = r_n T + r_n^2 T^2. \tag{A.11}$$

One readily verifies from (A.8) and (A.11) that

$$\limsup_{n \rightarrow \infty} \frac{1}{n^4} \mathbb{E}[\bar{\mathcal{E}}^n(T)]^2 = \bar{\Lambda}^2 \cdot (b-a)^2 T^2. \tag{A.12}$$

Now (A.1) follows from (A.9), (A.10) and (A.12). The proof is complete. ■

APPENDIX B. PROOF OF PROPOSITION 5.1

It is the goal of this appendix to establish the tightness of $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$, that is, to prove Proposition 5.1. The approach is to first show the tightness of $\{\zeta^{n,+} : n \geq 1\}$ and $\{\zeta^{n,-} : n \geq 1\}$ individually, and then obtain the joint tightness by proving the limit points of $\{\zeta^{n,+} : n \geq 1\}$ and $\{\zeta^{n,-} : n \geq 1\}$ are concentrated on the set of continuous paths. Throughout the proofs in this section, we use a generic constant C , which may vary from line to line, but C is independent of n .

We start with several auxiliary lemmas. The next lemma says that in order to establish the tightness of $\{\zeta^{n,+} : n \geq 1\}$ and $\{\zeta^{n,-} : n \geq 1\}$, it suffices to show that, for every $f \in C[0, 1]$, the sequences of real-valued processes $\{\langle \zeta^{n,+}, f \rangle : n \geq 1\}$ and $\{\langle \zeta^{n,-}, f \rangle : n \geq 1\}$ are both tight in $\mathbb{D}([0, T]; \mathbb{R})$. This lemma can be found in [18, Chapter 4, Proposition 1.7].

LEMMA B.1: A family of non-negative measure-valued processes $\{\nu^n : n \geq 1\}$ is tight in $\mathbb{D}([0, T], \mathcal{M}^+[0, 1])$ if $\{(\nu^n, f) : n \geq 1\}$ is tight in $\mathbb{D}([0, T]; \mathbb{R})$ for every $f \in C[0, 1]$.

The next lemma is Proposition VI.3.26 in [17].

LEMMA B.2: For fixed $T > 0$, let $\{\mathbb{X}^n : n \geq 1\}$ be a sequence of processes that take values in $\mathbb{D}([0, T], \mathbb{R}^d)$ equipped with Skorokhod J_1 topology. For each $n \geq 1$, $\{\mathbb{X}^n\}$ is adapted to a filtration $(\mathcal{F}_t^n)_{t \in [0, T]}$. Then $\{\mathbb{X}^n : n \geq 1\}$ is C -tight in $\mathbb{D}([0, T], \mathbb{R}^d)$ if and only if the following two conditions hold:

(a) For any $t \in [0, T]$ and $\epsilon > 0$, there exists a compact set $K(t, \epsilon) \subset \mathbb{R}^d$ such that

$$\inf_n \mathbb{P}(\mathbb{X}_t^n \in K(t, \epsilon)) > 1 - \epsilon. \tag{B.1}$$

(b) For every $\epsilon > 0$,

$$\lim_{\sigma \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\mathbb{X}_t^n - \mathbb{X}_s^n| > \epsilon \right) = 0. \tag{B.2}$$

For Lemma B.2 and its applications in this paper, we always take filtration $(\mathcal{F}_t^n)_{t \in [0, T]}$ as the natural filtration generated by $\{\mathcal{X}^n(nt) : t \in [0, T]\}$ for fixed $n \geq 1$. That is,

$$\mathcal{F}_t^n \triangleq \sigma(\mathcal{X}^n(ns), s \leq t). \tag{B.3}$$

The next lemma is Corollary VI.3.33 in [17].

LEMMA B.3: Let $\{\mathbb{X}^n : n \geq 1\}$ and $\{\mathbb{Y}^n : n \geq 1\}$ be two C -tight sequences of processes in $\mathbb{D}([0, T], \mathbb{R}^d)$. Then $\{\mathbb{X}^n + \mathbb{Y}^n : n \geq 1\}$ is C -tight in $\mathbb{D}([0, T], \mathbb{R}^d)$ and $\{(\mathbb{X}^n, \mathbb{Y}^n) : n \geq 1\}$ is C -tight in $\mathbb{D}([0, T], \mathbb{R}^{2d})$.

The next lemma states the C -tightness of the sequence of real-valued processes $\{\langle \zeta^n, f \rangle : n \geq 1\}$ for each $f \in C([0, 1])$. The proof is lengthy so we defer it to the end of this appendix. Recall from (3.3) that we have for each $n \geq 1$ and $t \in [0, T]$,

$$\langle \zeta_t^n, f \rangle = \frac{1}{n^2} \sum_{i=1}^n \mathcal{X}_i^n(nt) \cdot f\left(\frac{i}{n}\right).$$

LEMMA B.4: Fix any $f \in C([0, 1])$. The sequence of real-valued stochastic processes $\{\langle \zeta^n, f \rangle : n \geq 1\}$ is C -tight in $\mathbb{D}([0, T], \mathbb{R})$.

With Lemmas B.1–B.4 at our disposal, we are ready to prove Proposition 5.1.

PROOF OF PROPOSITION 5.1: We apply Lemma B.2 to establish the C -tightness of $\{\langle \zeta^{n,+}, f \rangle : n \geq 1\}$ in $\mathbb{D}([0, T], \mathbb{R})$ for fixed $f \in C([0, 1])$, where

$$\langle \zeta_t^{n,+}, f \rangle = \frac{1}{n^2} \sum_{i=1}^n \left[\mathcal{X}_i^n(nt) \right]^+ \cdot f\left(\frac{i}{n}\right). \tag{B.4}$$

The C -tightness of $\{\langle \zeta^{n,-}, f \rangle : n \geq 1\}$ follows using a similar argument. The C -tightness of $\{(\langle \zeta^{n,+}, f \rangle, \langle \zeta^{n,-}, f \rangle) : n \geq 1\}$ immediately follows after applying Lemma B.3.

We first show that $\{\langle \zeta_t^{n,+}, f \rangle : n \geq 1\}$ satisfies part (a) of Lemma B.2. We rely on the family of functions $f_\gamma \in C[0, 1]$ introduced in (5.1). Note that for fixed $t \in [0, T]$,

$$\begin{aligned} & \sup_n \mathbb{P}(|\langle \zeta_t^{n,+}, f \rangle| > L) \\ & \leq \sup_n \mathbb{P}\left(|\langle \zeta_t^{n,+}, f \rangle - \langle \zeta_t^n, f_\gamma \rangle| > \frac{L}{2}\right) + \sup_n \mathbb{P}\left(|\langle \zeta_t^n, f_\gamma \rangle| > \frac{L}{2}\right) \\ & \leq \frac{2}{L} \sup_n (\tilde{a}_n + C\gamma) + \sup_n \mathbb{P}\left(|\langle \zeta_t^n, f_\gamma \rangle| > \frac{L}{2}\right), \end{aligned}$$

where we use Markov inequality and (5.3) in the last inequality. Since $\{\tilde{a}_n\}$ is a bounded sequence, we can choose L large such that $\frac{2}{L} \sup_n (\tilde{a}_n + C\gamma)$ is arbitrarily small. In addition, for fixed γ and t , $\{\langle \zeta_t^n, f_\gamma \rangle : n \geq 1\}$ is a tight sequence by Lemma B.4. Thus, we can pick L large such that $\sup_n \mathbb{P}(|\langle \zeta_t^{n,+}, f \rangle| > L)$ is also arbitrarily small. These two facts imply that $\{\langle \zeta_t^{n,+}, f \rangle : n \geq 1\}$ is a tight sequence for fixed $t \in [0, T]$.

We next verify that $\{\langle \zeta_t^{n,+}, f \rangle : n \geq 1\}$ also satisfies part (b) of Lemma B.2. Clearly from (B.4) we have for fixed $n \geq 1$, the process $\langle \zeta_t^{n,+}, f \rangle$ is adapted to the filtration \mathcal{F}^n given in (B.3). Given a real number $c > 0$, we obtain from (5.3) that for n large

$$\begin{aligned} & \mathbb{P}\left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\langle \zeta_t^{n,+}, f \rangle - \langle \zeta_s^{n,+}, f \rangle| > c\right) \\ & \leq 2 \cdot \mathbb{P}\left(\sup_{0 \leq t \leq T} |\langle \zeta_t^{n,+}, f \rangle - \langle \zeta_t^n, f_\gamma \rangle| > \frac{c}{3}\right) + \mathbb{P}\left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\langle \zeta_t^n, f_\gamma \rangle - \langle \zeta_s^n, f_\gamma \rangle| > \frac{c}{3}\right) \\ & \leq 2 \cdot \frac{3}{c} (\tilde{a}_n + C\gamma) + \mathbb{P}\left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\langle \zeta_t^n, f_\gamma \rangle - \langle \zeta_s^n, f_\gamma \rangle| > \frac{c}{3}\right). \end{aligned} \tag{B.5}$$

Since for fixed $\gamma > 0$, the process $\langle \zeta_t^n, f_\gamma \rangle$ is adapted to \mathcal{F}^n , we infer from Lemma B.4 that

$$\lim_{\sigma \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\langle \zeta_t^n, f_\gamma \rangle - \langle \zeta_s^n, f_\gamma \rangle| > \frac{c}{3}\right) = 0.$$

In conjunction with (B.5), this implies for fixed $\gamma \in (0, p)$ and $c > 0$,

$$\lim_{\sigma \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\langle \zeta_t^{n,+}, f \rangle - \langle \zeta_s^{n,+}, f \rangle| > c\right) \leq \frac{6C\gamma}{c}.$$

Let $\gamma \rightarrow 0+$, we obtain

$$\lim_{\sigma \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\langle \zeta_t^{n,+}, f \rangle - \langle \zeta_s^{n,+}, f \rangle| > c\right) = 0.$$

Given the C -tightness of $\{\langle \zeta_t^{n,+}, f \rangle : n \geq 1\}$ and $\{\langle \zeta_t^{n,-}, f \rangle : n \geq 1\}$ for any fixed function $f \in C[0, 1]$, we first deduce from Lemma B.1 that both $\{\zeta_t^{n,+} : n \geq 1\}$ and $\{\zeta_t^{n,-} : n \geq 1\}$ are tight in $\mathbb{D}([0, T], \mathcal{M}^+[0, 1])$. We next argue that the limit points of $\{\zeta_t^{n,+} : n \geq 1\}$ and $\{\zeta_t^{n,-} : n \geq 1\}$ have continuous sample paths. Suppose ζ^+ is a limit point of $\{\zeta_t^{n,+} : n \geq 1\}$, then we readily verify from continuous mapping theorem that $\langle \zeta^+, f \rangle$ is a limit point of $\{\langle \zeta_t^{n,+}, f \rangle : n \geq 1\}$ for any function $f \in C[0, 1]$. Since all the limit points of the C -tight sequence $\{\langle \zeta_t^{n,+}, f \rangle : n \geq 1\}$ are concentrated on the set of continuous paths, we deduce that for almost every ω we have: for fixed $t \in [0, T]$,

$$\lim_{s \rightarrow t} d_+ \left(\zeta_s^+(\omega), \zeta_t^+(\omega)\right) = \lim_{s \rightarrow t} \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|\langle \zeta_s^+(\omega), \phi_k \rangle - \langle \zeta_t^+(\omega), \phi_k \rangle|}{1 + |\langle \zeta_s^+(\omega), \phi_k \rangle - \langle \zeta_t^+(\omega), \phi_k \rangle|} = 0, \tag{B.6}$$

where $\{\phi_k\}$ is a dense subset of $C[0, 1]$ and the distant measure d_+ is given in (5.5). Thus, the limit point ζ^+ has continuous sample paths. Similarly, almost every path of each limit point of

$\{\zeta^{n,-} : n \geq 1\}$ is continuous. Now the tightness of $\{(\zeta^{n,+}, \zeta^{n,-}) : n \geq 1\}$ in $\mathbb{D}([0, T], \overline{\mathcal{M}}([0, 1]))$ follows from Lemma 3.5 and Corollary 3.6 in [2]. ■

PROOF OF LEMMA B.4: We verify that the conditions (a) and (b) in Lemma B.2 are satisfied. For notational conveniences, we write for a fixed function $f \in C[0, 1]$ and for $t \geq 0$

$$\mathcal{Y}_t^n \triangleq \langle \zeta_t^n, f \rangle = \frac{1}{n^2} \sum_{i=1}^n \mathcal{X}_i^n(nt) \cdot f\left(\frac{i}{n}\right). \tag{B.7}$$

We first prove that \mathcal{Y}^n satisfies condition (a) from Lemma B.2. Since $\sup_{x \in [0, 1]} |f(x)| \leq C$ for some constant C by the continuity of f , we deduce from (B.7) that

$$|\mathcal{Y}_t^n| \leq \frac{1}{n^2} \sum_{i=1}^n |\mathcal{X}_i^n(nt)| \cdot |f(\frac{i}{n})| \leq \frac{C}{n^2} \sum_{i=1}^n |\mathcal{X}_i^n(nt)|.$$

We now deduce from Lemma A.1 that

$$\sup_n \mathbb{E} \left[\sup_{0 \leq t \leq T} |\mathcal{Y}_t^n|^2 \right] \leq C. \tag{B.8}$$

An application of Markov’s inequality immediately yields that \mathcal{Y}_t^n satisfies (B.1).

Next we show that \mathcal{Y}^n also satisfies condition (b) in Lemma B.2. Note that

$$\mathcal{Y}_t^n = F(\mathcal{X}^n(nt)) = \mathcal{Y}_0^n + \int_0^{nt} \mathcal{L}_n F(\mathcal{X}^n(s)) ds + \mathbb{M}_t^n, \tag{B.9}$$

where \mathcal{L}_n is the operator given in (2.5), \mathbb{M}^n is a (local) martingale, and the function F is defined in (5.11). Given $\epsilon > 0, \sigma > 0$, we deduce from (B.9) that

$$\begin{aligned} & \mathbb{P} \left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\mathcal{Y}_t^n - \mathcal{Y}_s^n| > \epsilon \right) \\ & \leq \mathbb{P} \left(\sup_{0 \leq t-s \leq \sigma, 0 \leq s, t \leq T} \left| \int_{ns}^{nt} \mathcal{L}_n F(\mathcal{X}^n(u)) du \right| > \frac{\epsilon}{2} \right) \\ & \quad + \mathbb{P} \left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\mathbb{M}_t^n - \mathbb{M}_s^n| > \frac{\epsilon}{2} \right). \end{aligned} \tag{B.10}$$

The strategy is to show for large n , there is a constant C independent of n such that

$$\mathbb{E} \left[\sup_{0 \leq t-s \leq \sigma, 0 \leq s, t \leq T} \left| \int_{ns}^{nt} \mathcal{L}_n F(\mathcal{X}^n(u)) du \right| \right] \leq C\sigma, \tag{B.11}$$

$$\mathbb{E} \left[\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\mathbb{M}_t^n - \mathbb{M}_s^n|^2 \right] \leq \frac{CT}{n^2}. \tag{B.12}$$

Using Markov’s inequality and Chebyshev’s inequality, and (B.10), we obtain

$$\lim_{\sigma \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |\mathcal{Y}_t^n - \mathcal{Y}_s^n| > \epsilon \right) = 0.$$

The rest of the proof focuses on establishing (B.11) and (B.12). We start with proving (B.11). For fixed n , we deduce from (5.11) and (2.5) that

$$\begin{aligned} \mathcal{L}_n F(\mathcal{X}^n(u)) &= \sum_{k < p_A^n(u)} \left[\frac{1}{n^2} f\left(\frac{k}{n}\right) \Theta_B^n(p_A^n(u) - k) |\mathcal{X}_k^n(u)| - \frac{1}{n^2} f\left(\frac{k}{n}\right) \Lambda_B^n(p_A^n(u) - k) \right] \\ &+ \sum_{k > p_B^n(u)} \left[\frac{1}{n^2} f\left(\frac{k}{n}\right) \Lambda_A^n(k - p_B^n(u)) - \frac{1}{n^2} f\left(\frac{k}{n}\right) \Theta_A^n(k - p_B^n(u)) |\mathcal{X}_k^n(u)| \right] \\ &+ \frac{1}{n^2} \left(f\left(\frac{p_B^n(u)}{n}\right) \Upsilon_A^n - f\left(\frac{p_A^n(u)}{n}\right) \Upsilon_B^n \right). \end{aligned} \tag{B.13}$$

Since f is bounded on $[0, 1]$, we deduce from Assumption 2.2 that

$$|\mathcal{L}_n F(\mathcal{X}^n(u))| \leq \frac{C}{n^3} \sum_{k=1}^n |\mathcal{X}_k^n(u)| + \frac{C}{n}. \tag{B.14}$$

This immediately yields that

$$\mathbb{E} \left[\sup_{0 \leq t-s \leq \sigma, 0 \leq s, t \leq T} \int_{ns}^{nt} |\mathcal{L}_n F(\mathcal{X}^n(u))| du \right] \leq \frac{C}{n^2} \mathbb{E} \left[\sup_{0 \leq s \leq T-\sigma} \int_s^{s+\sigma} \sum_{k=1}^n |\mathcal{X}_k^n(nu)| du \right] + C\sigma.$$

On combining (A.3), we obtain (B.11).

We next show (B.12). To this end, we define for each $u \geq 0$,

$$\eta^n(u) = \mathcal{L}_n F^2(\mathcal{X}^n(u)) - 2F(\mathcal{X}^n(u)) \cdot \mathcal{L}_n F(\mathcal{X}^n(u)),$$

where \mathcal{L}_n is given in (2.5) and F is the linear function given in (5.11). One checks that

$$\begin{aligned} \eta^n(u) &= \sum_{k < p_A^n(u)} \left[\frac{1}{n^4} f\left(\frac{k}{n}\right)^2 \cdot \Theta_B^n(p_A^n(u) - k) \cdot |\mathcal{X}_k^n(u)| + \frac{1}{n^4} f\left(\frac{k}{n}\right)^2 \cdot \Lambda_B^n(p_A^n(u) - k) \right] \\ &+ \sum_{k > p_B^n(u)} \left[\frac{1}{n^4} f\left(\frac{k}{n}\right)^2 \cdot \Lambda_A^n(k - p_B^n(u)) + \frac{1}{n^4} f\left(\frac{k}{n}\right)^2 \cdot \Theta_A^n(k - p_B^n(u)) \cdot |\mathcal{X}_k^n(u)| \right] \\ &+ \frac{1}{n^4} \left(f\left(\frac{p_A^n(u)}{n}\right)^2 \cdot \Upsilon_B^n + f\left(\frac{p_B^n(u)}{n}\right)^2 \cdot \Upsilon_A^n \right). \end{aligned}$$

Now for each fixed n , one verifies that $\{(M_t^n)^2 - \int_0^{nt} \eta^n(u) du : t \geq 0\}$ is a local martingale with respect to the filtration \mathcal{F}^n in (B.3) (see Lemma 5.1 in Appendix 1 of [18]). Suppose $\{(M_t^n)^2 - \int_0^{nt} \eta^n(u) du : t \geq 0\}$ and $\{M_t^n : t \geq 0\}$ are indeed \mathcal{F}^n -martingales. Then Doob’s maximal inequality for martingales yields that

$$\mathbb{E} \left[\sup_{|s-t| \leq \sigma, 0 \leq s, t \leq T} |M_t^n - M_s^n|^2 \right] \leq 4\mathbb{E} \left[\sup_{0 \leq t \leq T} |M_t^n|^2 \right] \leq 16\mathbb{E}(|M_T^n|^2) = \mathbb{E} \left[\int_0^{nT} \eta^n(u) du \right] \tag{B.15}$$

Hence to establish (B.12), we proceed to bound $\eta^n(u)$. Using Assumption 2.2 and the fact that f is bounded on $[0, 1]$, there is a constant $C > 0$ such that

$$0 \leq \eta^n(u) \leq \frac{1}{n^4} \left[\frac{C}{n} \sum_{k=1}^n |\mathcal{X}_k^n(u)| + Cn \right]. \tag{B.16}$$

Upon combining (A.3) and applying the change of variable formula we find for n large

$$\mathbb{E} \left[\int_0^{nT} \eta^n(u) du \right] \leq \frac{CT}{n^2}. \tag{B.17}$$

We then obtain (B.12) from (B.15).

It remains to show that the two local martingales $\{(M_t^n)^2 - \int_0^{nt} \eta^n(u) du : t \geq 0\}$ and $\{M_t^n : t \geq 0\}$ are indeed \mathcal{F}^n -martingale for each fixed n . By [29, Theorem 51], it suffices to show for every $t \leq T$,

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |M_s^n|^2 \right] < \infty, \tag{B.18}$$

$$\mathbb{E} \left[\int_0^{nt} \eta^n(u) du \right] < \infty. \tag{B.19}$$

Inequality (B.19) directly follows from (B.17). To prove (B.18), we use (B.9), which implies

$$\sup_{0 \leq s \leq t} |M_s^n|^2 \leq 3 \sup_{0 \leq s \leq t} (\mathcal{Y}_s^n)^2 + 3(\mathcal{Y}_0^n)^2 + 3 \left(\int_0^{nt} \left| \mathcal{L}_n F(\mathcal{X}^n(s)) \right| ds \right)^2$$

Inequality (B.18) then follows from (B.8), (B.14) and (A.1). The proof is thus complete. ■

APPENDIX C. PROOF OF PROPOSITION 5.2

In this section, we prove Proposition 5.2. We rely on the representation in (5.12), that is,

$$\begin{aligned} \langle \zeta_t^n, f \rangle &= \frac{1}{n^2} \sum_{i=1}^n \mathcal{X}_i^n(nt) \cdot f\left(\frac{i}{n}\right) = F(\mathcal{X}^n(nt)) \\ &= \langle \zeta_0^n, f \rangle + n \int_0^t \mathcal{L}_n F(\mathcal{X}^n(ns)) ds + M_t^n. \end{aligned} \tag{C.1}$$

Here \mathcal{L}_n is operator given in (2.5), M^n is a martingale and F is given in (5.11). We fix $f \in C([0, 1])$ throughout this section. It is clear from (2.6) that

$$\lim_{n \rightarrow \infty} \langle \zeta_0^n, f \rangle = \int_0^1 f(x) \varrho(x) dx = \langle \zeta_0, f \rangle, \tag{C.2}$$

where ζ_0 is a deterministic signed measure with density $\zeta_0(dx) = \varrho(x)dx$ for $x \in [0, 1]$.

To prove Proposition 5.2, we now introduce two auxiliary lemmas. The next lemma implies that the martingale term in (C.1) vanishes (converges weakly to zero) as $n \rightarrow \infty$. The proof directly follows from (B.15) and (B.17).

LEMMA C.1: *For $T > 0$, we have*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{0 \leq t \leq T} |M_t^n|^2 \right] = 0. \tag{C.3}$$

The next lemma concerns the weak convergence of the “scaled” generator. The proof is lengthy and thus deferred to the end of this appendix.

LEMMA C.2: *For the subsequence $\{n_k : k = 1, 2, \dots\}$ in (5.7) we have for fixed $f \in C([0, 1])$*

$$\sup_{0 \leq s \leq T} \left| n_k \cdot \mathcal{L}_{n_k} F(\mathcal{X}^{n_k}(n_k s)) - \langle \nu, f \rangle + \langle \zeta_s, \mathcal{A}_\Theta f \rangle \right| \Rightarrow 0 \quad \text{as } n_k \rightarrow \infty, \tag{C.4}$$

where ν is given in (5.13), \mathcal{A}_Θ is given in (5.14) and F is given in (5.11).

We are now ready to prove Proposition 5.2.

PROOF OF PROPOSITION 5.2: Given (C.1)–(C.4) and that $\langle \zeta^{n_k}, f \rangle \Rightarrow \langle \zeta, f \rangle$ as $n_k \rightarrow \infty$, we deduce that the process ζ satisfies Eq. (5.15). The continuity of the path of $\langle \zeta, f \rangle$ directly follows from Lemma B.4.

Next we show that there is an unique measure-valued process satisfying (5.15). It is evident that ζ is deterministic since there is no source of randomness in Eq. (5.15). Suppose for fixed $t \in [0, T]$, ζ_t is absolutely continuous with respect to Lebesgue measure and it has a bounded density function φ such that $\zeta_t(dx) = \varphi(x, t)dx$. After multiplying a smooth test function $g(t)$ on both sides of Eq. (5.15), integrating with respect to t on $[0, T]$, substituting (5.13) and (5.14), and noting that f can be an arbitrary continuous function on $[0, 1]$, we readily verify that for almost every $x \in (p, 1]$,

$$\int_0^T \varphi(x, t)g(t)dt = \varrho(x) \int_0^T g(t)dt + \int_0^T g(t) \int_0^t [\Lambda_A(x - p) - \Theta_A(x - p)\varphi(x, s)] ds dt.$$

That is, for fixed $x \in (p, 1]$, $\varphi(x, \cdot)$ is a weak solution for the ODE given in (3.6) and (3.7), and thus $\varphi(x, t)$ is measurable in the second argument t . On the other hand, it is clear that there is an unique classical solution to the ODE (3.6) and (3.7) for fixed $x \in (p, 1]$. Therefore, we deduce from the equivalence of classical solution and weak solution for ODE (see, e.g., [32, Chapter 1, Lemma 1.3]) that the density function $\varphi(x, \cdot)$ is the unique classical solution of (3.6) and (3.7) when $x > p$. A similar argument yields that $\varphi(x, \cdot)$ is the unique classical solution for (3.6) and (3.8) when $x < p$. As a consequence, the solution for (5.15) is unique (we allow ourselves the ability to modify $\varphi(x, t)$ on a set of measure zero including the point $x = p$. The modification is independent of t).

The rest of the proof focuses on showing that ζ_t is absolutely continuous and that it has a bounded density function. We prove that there exists some constant C_T depending on T such that for all $f \in C([0, 1])$ and $t \in [0, T]$,

$$|\langle \zeta_t, f \rangle| \leq C_T \int_0^1 |f(x)|dx.$$

To this end, we first deduce from Eq. (5.10) and the continuous mapping theorem that for any $f \in C([0, 1])$ and $t \in [0, T]$,

$$|\langle \zeta_t^{n_k}, f \rangle| \Rightarrow |\langle \zeta_t, f \rangle| \quad \text{as } n_k \rightarrow \infty.$$

Applying Theorem 3.4 in [4] yields

$$\mathbb{E}|\langle \zeta_t, f \rangle| \leq \liminf_{n_k \rightarrow \infty} \mathbb{E}|\langle \zeta_t^{n_k}, f \rangle|. \tag{C.5}$$

We next focus on bounding $\mathbb{E}|\langle \zeta_t^{n_k}, f \rangle|$. Recall the limit order arrival process A_i^n is pathwisely upper bounded by a Poisson process \bar{A}_i^n with rate $\bar{\Lambda}$ (see (A.7)). Hence,

$$\begin{aligned} |\langle \zeta_t^n, f \rangle| &= \left| \frac{1}{n^2} \sum_{i=1}^n \mathcal{X}_i^n(nt) \cdot f\left(\frac{i}{n}\right) \right| \\ &\leq \frac{1}{n^2} \sum_{i=1}^n |\mathcal{X}_i^n(0)| \cdot \left| f\left(\frac{i}{n}\right) \right| + \left| \frac{1}{n^2} \sum_{i=1}^n A_i^n(nt) \cdot f\left(\frac{i}{n}\right) \right| \\ &\leq \frac{1}{n} \sum_{i=1}^n |\varrho\left(\frac{i}{n}\right)| \cdot \left| f\left(\frac{i}{n}\right) \right| + \frac{1}{n^2} \sum_{i=1}^n \bar{A}_i^n(nt) \cdot \left| f\left(\frac{i}{n}\right) \right|, \end{aligned}$$

which implies that

$$\mathbb{E}|\langle \zeta_t^{n_k}, f \rangle| \leq \frac{1}{n_k} \sum_{i=1}^{n_k} \left| \varrho\left(\frac{i}{n_k}\right) \right| \cdot \left| f\left(\frac{i}{n_k}\right) \right| + \bar{\Lambda}t \cdot \frac{1}{n_k} \sum_{i=1}^{n_k} \left| f\left(\frac{i}{n_k}\right) \right|.$$

In conjunction with (C.5) and the fact that ζ_t is a deterministic measure, we deduce

$$|\langle \zeta_t, f \rangle| \leq \int_0^1 |\varrho(x)| \cdot |f(x)| dx + \bar{\Lambda}t \cdot \int_0^1 |f(x)| dx, \tag{C.6}$$

after letting $n_k \rightarrow \infty$. Thus, we have all $f \in C([0, 1])$ and $t \in [0, T]$,

$$|\langle \zeta_t, f \rangle| \leq C_T \int_0^1 |f(x)| dx, \tag{C.7}$$

where $C_T = \max_{x \in [0, 1]} |\varrho(x)| + \bar{\Lambda}T$. Since continuous functions are dense in the space of Lebesgue-integrable functions $L^1([0, 1])$, we deduce from the bounded linear transformation theorem that Eq. (C.7) holds for all $f \in L^1([0, 1])$. We then conclude from [11, Theorem 15.6] that ζ_t is a finite measure, absolutely continuous with respect to Lebesgue measure, and its density function $\varphi(x, t)$ is bounded by C_T uniformly with respect to $x \in [0, 1]$ and $t \in [0, T]$. The proof is complete. ■

PROOF OF LEMMA C.2: Throughout the proof, we fix $f \in C([0, 1])$ and we use $\{n : n \geq 1\}$ instead of its subsequence $\{n_k : k \geq 1\}$ for notational simplicity. We also use a generic constant C which may vary from line to line but C is independent of n .

We first show for $f \in C[0, 1]$,

$$\sup_{0 \leq s \leq T} \left| n \cdot \sum_{k > p_B^n(ns)} \left[\frac{1}{n^2} f\left(\frac{k}{n}\right) \cdot \Lambda_A^n(k - p_B^n(ns)) \right] - \int_{\mathbf{p}}^1 f(x) \Lambda_A(x - \mathbf{p}) dx \right| \Rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{C.8}$$

To this end, we define for $z \in [0, 1]$ and $n \geq 1$

$$h_n(z) = \sum_{\frac{k}{n} > z} \left[\frac{1}{n} f\left(\frac{k}{n}\right) \Lambda_A\left(\frac{k}{n} - z\right) \right], \tag{C.9}$$

$$h(z) = \int_z^1 f(x) \Lambda_A(x - z) dx. \tag{C.10}$$

Applying Assumption 2.2, we find that (C.8) is equivalent to

$$\sup_{0 \leq s \leq T} \left| h_n\left(\frac{p_B^n(ns)}{n}\right) - h(\mathbf{p}) \right| \Rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{C.11}$$

Part (a) of Theorem 3.1 implies that when $n \rightarrow \infty$, we have

$$\frac{p_B^n(n \cdot)}{n} \Rightarrow \mathbf{p} \quad \text{in } \mathbb{D}([0, T], \mathbb{R}).$$

In addition, since f and Λ_A are uniformly continuous and bounded on $[0, 1]$, it is straightforward to verify that the function h is continuous, and the sequence $\{h_n : n \geq 1\}$ converges uniformly to h . In particular, we obtain for any sequence of real numbers $\{z_n : n \geq 1\}$,

$$\lim_{n \rightarrow \infty} h_n(z_n) = h(\mathbf{p}) \quad \text{if } \lim_{n \rightarrow \infty} z_n = \mathbf{p} \in (0, 1). \tag{C.12}$$

Then generalized continuous mapping theorem [33, Theorem 3.4.4] yields that as $n \rightarrow \infty$

$$h_n\left(\frac{p_B^n(n \cdot)}{n}\right) \Rightarrow h(\mathbf{p}) \quad \text{in } \mathbb{D}([0, T], \mathbb{R}).$$

Since the Skorohod J_1 topology relativized to $C([0, T], \mathbb{R})$ of continuous functions is equivalent to the uniform topology [4, Section 12], we obtain (C.11) and (C.8).

Applying a similar argument, we find

$$\sup_{0 \leq s \leq T} \left| \sum_{k < p_A^n(ns)} \left[\frac{1}{n} f \left(\frac{k}{n} \right) \cdot \Lambda_B^n(p_A^n(ns) - k) \right] - \int_0^p f(x) \Lambda_B(p - x) dx \right| \Rightarrow 0 \quad \text{as } n \rightarrow \infty. \tag{C.13}$$

We next prove that when $n \rightarrow \infty$,

$$\sup_{0 \leq s \leq T} \left| \sum_{k > p_B^n(ns)} \left[\frac{1}{n} f \left(\frac{k}{n} \right) \cdot \Theta_A^n(k - p_B^n(ns)) \cdot \mathcal{X}_k^n(ns) \right] - \int_p^1 f(x) \Theta_A(x - p) d\zeta_s(x) \right| \Rightarrow 0. \tag{C.14}$$

It suffices to show

$$\sup_{0 \leq s \leq T} \left| \int_p^1 f(x) \Theta_A(x - p) d\zeta_s^n(x) - \int_p^1 f(x) \Theta_A(x - p) d\zeta_s(x) \right| \Rightarrow 0, \tag{C.15}$$

and

$$\sup_{0 \leq s \leq T} \left| \sum_{k > p_B^n(ns)} \left[\frac{1}{n} f \left(\frac{k}{n} \right) \Theta_A^n(k - p_B^n(ns)) \cdot \mathcal{X}_k^n(ns) \right] - \int_p^1 f(x) \Theta_A(x - p) d\zeta_s^n(x) \right| \Rightarrow 0. \tag{C.16}$$

First, we prove (C.15). We provide a sketch. Define a function G by setting

$$G(x) = \begin{cases} f(x) \Theta_A(x - p) & \text{if } p < x \leq 1, \\ 0 & \text{if } 0 \leq x \leq p. \end{cases} \tag{C.17}$$

If G is continuous on $[0, 1]$, (C.15) readily follows from (5.10) and the equivalence of Skorohod J_1 topology and the uniform topology on the space of continuous functions. If G is not continuous, one can similarly prove (C.15) by constructing a family of continuous functions $\{G_\epsilon : \epsilon > 0\}$ such that for each small $\epsilon > 0$, G and G_ϵ are equal except on a small interval $(p - \epsilon, p]$. We omit further details.

Next, we prove (C.16). We also provide a sketch. Using Assumption 2.2 and the definition of ζ^n in (3.3), it is equivalent to show as $n \rightarrow \infty$

$$\sup_{0 \leq s \leq T} \left| \sum_{\frac{k}{n} > \frac{p_B^n(ns)}{n}} \left[\frac{1}{n^2} f \left(\frac{k}{n} \right) \Theta_A \left(\frac{k}{n} - \frac{p_B^n(ns)}{n} \right) \cdot \mathcal{X}_k^n(ns) \right] - \sum_{\frac{k}{n} > p} \left[\frac{1}{n^2} f \left(\frac{k}{n} \right) \Theta_A \left(\frac{k}{n} - p \right) \cdot \mathcal{X}_k^n(ns) \right] \right| \Rightarrow 0. \tag{C.18}$$

This can be readily verified from (3.4), (3.5) and Lemma A.1. We omit further details.

Similarly we can show as $n \rightarrow \infty$,

$$\sup_{0 \leq s \leq T} \left| \sum_{k < p_A^n(ns)} \left[\frac{1}{n} f \left(\frac{k}{n} \right) \Theta_B^n(p_A^n(ns) - x) \cdot \mathcal{X}_k^n(ns) \right] - \int_0^p f(x) \Theta_B(p - x) d\zeta_s(x) \right| \Rightarrow 0. \tag{C.19}$$

Finally, using Assumption 2.2 and the fact that f is bounded on $[0, 1]$, we obtain that there exists some constant C such that

$$\sup_{0 \leq s \leq T} n \left| \frac{1}{n^2} \left(f \left(\frac{p_B^n(u)}{n} \right) \Upsilon_A^n - f \left(\frac{p_A^n(u)}{n} \right) \Upsilon_B^n \right) \right| \leq \frac{C}{n^{1-\kappa}} \Rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{C.20}$$

On combining (C.8), (C.13), (C.14), (C.19), and (C.20), we obtain (C.4). ■