A multiscale model of high-frequency trading

Andrei Kirilenko\textsuperscript{a}, Richard B. Sowers\textsuperscript{b,\ast}, Xiangqian Meng\textsuperscript{c}

\textsuperscript{a}Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02142-1347, USA
E-mail: ak67@mit.edu

\textsuperscript{b}Department of Industrial and Enterprise Systems Engineering and Department of Mathematics, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA
E-mail: r-sowers@illinois.edu

\textsuperscript{c}E-mail: memorylemon@gmail.com

Abstract. We propose and study a stylization of high frequency trading (HFT). Our interest is an order book which consists of orders from slow liquidity traders and orders from high-frequency traders. We would like to frame a model which is amenable to the (seemingly natural) mathematical toolkit of separation of scales and which can be used to address some of the larger issues involved in HFT.

The main issue to which we address our model is volatility. An important question is how volatility is affected by HFT. In our stylized model, we show how HFT increases volatility, and can quantify this effect as a function of the parameters in our model and the separation of scales.

Keywords: High frequency trading, volatility

1. Introduction

One of the issues which is become increasingly important in finance is the effect of high-frequency trading (HFT). While high frequency traders (HFT-ers) claim that they provide liquidity, the volume they trade and the algorithmic nature of HFT can cause significant impact. The pre-eminent example of this, of course, is the Flash Crash of May 6th of 2011. That event highlighted the need for a thorough effort in modelling, evaluating, and controlling the effects of HFT.

We believe that quantitative modelling of HFT is an important issue. There is a concern that the Flash Crash and HFT has affected the stability of the markets to such an extent that investors are becoming wary of the market. A better understanding of HFT-related phenomena is clearly necessary, both due to its importance in the modern financial system, and also to provide a theoretical underpinning for market controls (viz. measures like circuit-breakers).

The finance community has already initiated a number of empirical studies of HFT. There is also a small (but increasing) number of papers from the mathematical finance community; see, for example, the works of Avellaneda and Stoikov (2008), Cont (2011), Cont and De Larrard (n.d.a), Cont and De Larrard (n.d.b), Cvitanic and Kirilenko (n.d.), Jarrow and Protter (2012), and Moallemi and Saglam (n.d.). We believe that in fact the mathematical community has a collection of tools which are uniquely adapted to several issues related to HFT. In contrast to the main thrusts of modern finance and economics, which often focus on identifying and understanding macroscopic market behaviors, HFT is intrinsically microscopic; small amounts of money are traded large numbers of times, and an important framework seems to be the mathematics of multiscale asymptotics, in which different scales of a system interact.
Our intent here is to develop and study a very stylized model of an order book which is populated by a community of HFT-ers and a community of liquidity traders. The HFT-ers operate at a much faster scale than the liquidity traders. They also place and cancel a large number of orders (which is indeed what happens; HFT can account for over half of the trades on a typical day (Oloffson & Gandel, 2009)). Ex post, we want to model the facts

- The dominant price dynamics are diffusive and are due to lower-frequency traders
- HFT-ers add and remove liquidity at a much faster rate than lower-frequency traders.

Although HFT-ers make money with a range of (highly proprietary) specific strategies, our aim is to model the macroscopic effect of HFT, and not any specific strategy. As such, our model will be designed to focus on the simplest incarnation of the phenomena we wish to study.

We are interested in focusing on a separation of scales. The mathematics of multiscale systems attempts to identify how a system with several scales can be approximated by a simpler system. HFT phenomena seems to be an area which is rife with issues of this nature. We in particular want to understand how a model for HFT-ers and a separation of scales leads to price changes and thus affects the volatility. While our model could probably not be characterized as ‘simple’, we feel it is the simplest one which captures the effects of interest. There are a number of phenomena and interactions which we have intentionally neglected; the value of our model (we believe) is that it provides a sample case where the entire transition from microstructure to macroscopic volatility can be explicitly calculated.

2. Price dynamics

The main focus of our efforts is price dynamics of a single asset. We want to start with a model for market microstructure which involves liquidity traders (LT-ers) and high-frequency traders, and we ultimately want to understand the macroscopic implications of this model. In fact, ‘price’ can refer either to the last traded price or the midpoint price. We will focus on the midpoint price; this essentially translates our interest into a study of liquidity. How big are the queues for the bid and ask orders (the Level I queues)? The midpoint price changes when either of these queues are emptied, either by execution (crossing the order book) or cancellation (thus the dynamics of the midpoint price are slightly simpler than that of trade prices; where cancellations need to be distinguished from executions). By focussing on the size of the order book, we can access a wide range of results about the asymptotic theory of queues (i.e., fluid limits and heavy traffic limits). These asymptotic results will play central roles in our analysis.

Our model has three time scales. On the macroscopic scale, the price fluctuations are of order 1. The mesoscopic scale, which we will control via a small parameter $\delta \in (0, 1)$, will correspond to the tick size of the stock. Liquidity traders operate on a time scale defined by $\delta$. High frequency traders operate on a scale faster than that of the liquidity traders, and we will characterize their time scale via another small parameter $\epsilon \in (0, 1)$. Informally, we want $\epsilon$ to be small compared to $\delta$. More precise relations between $\delta$ and $\epsilon$ will be given in Definitions 4.6 and 7.1. Informally, we want the high-frequency traders to operate in a regime which is faster than that of the liquidity traders, but slow enough that the dominant correction (i.e., the “next term”) to the liquidity-induced price dynamics is caused by the high-frequency trading (see Remark 4.4).

Our efforts are organized as follows. There are two major components in our efforts;

- modelling and analysis of the order book dynamics with the goal of characterizing the times (the “epochs”) at which the midpoint price changes; Sections 3–5
- translating the statistical description of these epochs into statements about volatility; Sections 7.

Again, our goal is to construct a simplified model where this entire transition can be rigorously justified. Our model is at best a pale approximation of reality.

The order book for the asset consists of a collection of bid orders and a collection of ask orders. We will start by focussing on only one of these queues; i.e., either the queue of bid orders or the queue of ask orders. In Section 3, we will model the behavior of the liquidity traders. Essentially, we want to model the order activity due to liquidity traders by techniques from fluid limits for queues. We want to arrange things so that the resulting price dynamics are diffusive. In Section 4, we model high-frequency traders. Orders appear (i.e., are added) and disappear (due to cancellation or execution) at a faster rate,
but are balanced in a way which does not affect the dominant time at which the queue empties. Roughly, this corresponds to a heavy-traffic perturbative effect. We then identify the resulting approximate diffusive behavior of the order queue, and in Subsection 5 we precisely quantify the effect of the high-frequency traders on the time at which the queue empties; i.e. the 'epochs'.

In Section 7 we return to a consideration of the combined order book and the joint behavior of the bid and ask queues. Our goal is to characterize how the statistical variation in the epochs affects volatility. The heuristic picture is natural. If the order queues empty earlier due to high-frequency trading, the price will have faster variations, and volatility should increase. If the order queues empty later, volatility should decrease. In our simplified model, we are able to precisely characterize these effects.

Our main result is Theorem 7.6, which quantifies how HFT-ers (in our model) affect volatility. The result is perturbative; in our model HFT-ers provide a small correction to the dominant effect of liquidity traders.

Some remarks about notation are in order. We will let \((\Omega, \mathcal{F}, \mathbb{P})\) be our underlying probability triple on which all random variables are defined. Since the underlying dynamics of our model are Markovian, we will heavily use the martingale characterization of Markov processes as developed in Ethier and Kurtz (1986). We believe that, given the multiple levels of our model, this formulation provides the best combination of clarity and readability.

An intrinsic component of our model is the Poisson process. We will use both fluid and heavy-traffic approximations of queues. To recall these results, suppose that \(N\) is a standard Poisson process. If we define

\[
V^\kappa_t \overset{\text{def}}{=} \sqrt{\kappa} \left\{ N \left( \frac{t}{\kappa} \right) - \frac{t}{\kappa} \right\}
\]

for all \(t \geq 0\) and \(\kappa > 0\), then \(V^\kappa\) tends to a Brownian motion as \(\kappa \downarrow 0\). Writing

\[
\kappa N \left( \frac{t}{\kappa} \right) \approx t + \sqrt{\kappa} V^\kappa_t,
\]

the fluid approximation of \(N\) is the (deterministic) first term and the heavy-traffic approximation is the (diffusive) second term.

There is a growing interest in modelling high frequency trading. Our work is heavily influenced by the works of Cont and De Larrard (n.d.a), Cont and De Larrard (n.d.b), and Cont et al. (2010), which developed and analyzed a rigorous model for the order book based on ideas from queueing theory. Our contributions are twofold. First, we consider an explicitly multiscale model where the order queue dynamics are driven by two types of traders (the liquidity traders and the high-frequency traders). This requires a more detailed analysis of the interaction between various scales. Driving the dynamics by these two scales then allows us to precisely characterize the effect of the smaller (microscopic; i.e., HFT-driven) scale, and to do so in a perturbative way. Our second contribution is stochasticity in the rate at which the high-frequency traders add and remove liquidity. In a way similar to stochastic volatility, the stochasticity in our rates allow a slightly richer collection of models; our interest is in modelling the combined effect of a large population of such high-frequency traders (see Remark 4.3). Theorem 7.6 precisely quantifies how HFT (in our model) produces an asymptotic correction to the asset volatility (in the asymptotic regime we study). Namely, the excess volatility in our model is proportional to

\[
\text{volatility of HFT arrival rate} \times \sqrt{\text{(rate at which LT'ers empty queue)}} \times (\text{tick size})^2 \div (\text{initial size of queue})^{3/2}. \tag{2}
\]

See also Remark 7.7.

We believe that scaling relationships can play an important role in understanding the roles market participants working at different speeds. Scaling relationships often give rise to invariants which allow one to structure data analysis. Other work in this area is due to Kyle and Obizhaeva (n.d.).

There are two major oversights in our work. Firstly, we consider a very simplified model for the structure of the initial condition of the bid and ask queues right after the epochs at which the midpoint price changes. For simplicity, we have assumed that the queues start at a fixed level. A more realistic picture would stem from incorporating the corresponding aspects (i.e., the model for the initial condition of the queues) of the works of Cont and De Larrard (n.d.a) and Cont and De Larrard (n.d.b) into our model. Given the complexities of our existing effort, we have decided to leave that for future work. A second oversight is that our model of the dynamics of liquidity traders
3. Liquidity traders

Let’s begin with liquidity traders (LT-er’s); i.e., let’s assume that there are no HFT-ers. Fix a small parameter $\delta \in (0, 1)$ which represents the tick size; the activity of the LT-ers will be written in terms of $\delta > 0$. The order queue will be defined by a jump Markov process which is governed by small but rapid additions and deletions from the order queue (i.e., rapid addition and subtraction of small amounts of liquidity). We want the various size parameters of the LT-ers to be organized so that the midpoint price moves (i.e., the queue empties) in a diffusive way.

Let’s fix a $q \in (0, \infty)$ and an exponent $\iota$ (which we will specify in a moment in (10)) and assume that the order queue starts with $q\delta$. To specify the dynamics of the order queue (due to the LT-ers), we need four parameters; two positive exponents $\nu_\iota$ and $\gamma_\iota$ and two rates $\lambda_\iota^+$ and $\lambda_\iota^-$ in $(0, \infty)$. The LT-ers add orders of size $\delta^\nu_\iota$ with rate $\delta^{-\gamma_\iota} \lambda_\iota^+$ (i.e., they add liquidity in increments of $\delta^\nu_\iota$ at rate $\delta^{-\gamma_\iota} \lambda_\iota^+$). The LT-ers cancel and execute orders of size $\delta^\nu_\iota$ at a combined rate of $\delta^{-\gamma_\iota} \lambda_\iota$ (i.e., they remove liquidity in increments of $\delta^\nu_\iota$ at rate $\delta^{-\gamma_\iota} \lambda_\iota^-$). In yet other words, the order queue decreases by $\delta^\nu_\iota$ at rate $\delta^{-\gamma_\iota} \lambda_\iota^-$. To proceed, we mathematically formalize our model of the order book when there are only liquidity traders (and under the assumption that the order book is a Markov process). For $f \in B(\mathbb{R})$, define

$$(A_\delta f)(x) = \frac{1}{\delta^{\nu_\iota}} \{f(x + \delta^\nu_\iota) - f(x)\} \lambda_\iota^+$$

$$+ \frac{1}{\delta^{\gamma_\iota}} \{f(x - \delta^\nu_\iota) - f(x)\} \lambda_\iota^- \quad x \in \mathbb{R} \quad (3)$$

For every $f \in B(\mathbb{R})$, we then have that

$$f(Q^\delta_t) = f(q\delta^\iota) + \int_{s=0}^{t} (A_\delta f)(Q^\delta_s) ds + M_t$$

where $M$ is a martingale with respect to the filtration

$$\mathcal{F}_t^L \triangleq \sigma(Q^\delta_s; 0 \leq s \leq t).$$

Define now

$$\tau^L_\delta \triangleq \inf \left\{ t > 0 : Q^\delta_t \leq 0 \right\} \quad (4)$$

this is the time at which the queue empties. We will model the behavior of the order queue as the process $\{(Q^\delta_{t\wedge \tau^L_\delta})^+ ; t \geq 0\}$. In other words, the order queue
evolves like $Q^{L,\delta}$ up to the time $\tau^{L,\delta}$ at which time it jumps to 0 and stays there (i.e., the queue empties).

Our model attempts to in some sense capture the ensemble average of a community of liquidity traders. Different traders will follow different strategies, and our model only seeks to capture the overall behavior of the collection of LT-ers. Any single LT-er would seek to optimize profit under a more refined model. Overall, however, a large number of traders using different models would in some sense average out. The collective behavior is what we hope to model.

We want to understand the behavior of the liquidity traders via a continuum approximation. Liquidity addition and removal due to the LT-ers is governed by

$$\delta^{\nu_{\epsilon}} \bar{N}_{L}^+ (\delta^{-\gamma_{\epsilon}} \lambda_{L}^{2} \ell t)$$

where $\bar{N}_{L}^+$ and $\bar{N}_{L}^-$ are two independent Poisson processes. From (1), we should expect that

$$Q_{t}^{L,\delta} \approx q \delta^{-\gamma_{\epsilon}} N_{L}^+ (\delta^{-\gamma_{\epsilon}} \lambda_{L}^{2} \ell) - \delta^{\nu_{\epsilon}} \bar{N}_{L}^-$$

$$\approx q \delta^{-\gamma_{\epsilon}} \ell t + \delta^{\nu_{\epsilon} - \gamma_{\epsilon}}/2 \{ \hat{V}_{L}^{+} t - \hat{V}_{L}^{-} t \}$$

where

$$\ell \overset{\text{def}}{=} \lambda_{L}^{-} - \lambda_{L}^{+}$$

and where $\hat{V}_{L}^{\pm}$ are two standard Brownian motions. Informally, $\ell$ is the rate at which the order queue empties due to liquidity trading. The dominant behavior of the order queue is thus

$$Q_{t}^{L,\delta} \approx q \delta^{-\gamma_{\epsilon}} \ell t$$

up until time $\tau^{L,\delta}$.

We want to arrange things so that the midpoint price changes diffusively due to the LT-ers. When the queue empties, the midpoint price will change by a tick; i.e., by $\delta$. If the approximation (7) is good enough, we need that

$$\ell > 0$$

i.e., $\lambda_{L}^{-} > \lambda_{L}^{+}$. Then $E[\tau^{L,\delta}] < \infty$ (see Lemma A.1; see also the works of Cont and De Larrard (n.d.b), Cont and De Larrard (n.d.a), and Cont et al. (2010) for some related asymptotics when (8) is relaxed). Rearranging (7), we should expect that if (8) holds, then

$$\tau^{L,\delta} \approx q \delta^{\epsilon} \overline{\delta^{\nu_{\epsilon} - \gamma_{\epsilon}} \ell t}.$$  

To get the midpoint price to change at a diffusive rate (an ex post modelling step), we need the queue to empty in a time scale of order $\delta^2$. We thus require

$$\ell = 2 + \nu_{\epsilon} - \gamma_{\epsilon}.$$  

Define

$$\hat{T} \overset{\text{def}}{=} \frac{q}{\ell},$$

using (10) in (9), we then expect that

$$\tau^{L,\delta} \approx \delta^{2} \hat{T}.$$  

We will think of the effect of LT-ers in the absence of HFT-ers’ as the ‘reference’ case. We will want to think of the effect of HFT-ers as a perturbation of this case. To make sure that we are working on solid footing, let’s rigorously prove (12). An added benefit is that we can develop several arguments which are indicative of the tools we will use for the HFT-ers. Let’s now enforce the assumption that

$$\nu_{\epsilon} > 2.$$  

With these asymptotics, (12) can be made precise.

**Lemma 3.1.** We have that $\lim_{\delta \searrow 0} \delta^{-2} \tau^{L,\delta} = \hat{T}$ in probability; i.e., for every $\kappa > 0$ we have that

$$\lim_{\delta \searrow 0} \mathbb{P} \left( \left| \frac{\tau^{L,\delta}}{\delta^2} - \hat{T} \right| \geq \kappa \right) = 0.$$

We will delay the proof until Appendix A.

**Remark 3.2.** There are a number of ways we could generalize our model of liquidity traders; we could, for example, make the increments or rates random. Since our interest lies more in the effect of high-frequency traders, our model of liquidity traders is as simple as possible.
4. High frequency traders

Let’s next add some high-frequency traders. As in the case of liquidity traders, we first of all model the rates and increments by which liquidity is added and removed due to HFT-ers. Adding this to our existing model of LT-ers, we will get a multiscale model for the order queue in which the size of the queue is affected by both liquidity traders and high-frequency traders. As in our model for liquidity traders, we will assume that the queue dynamics are Markovian. Here, we want liquidity addition and removal to be balanced, so that the LT-ers still play the dominant role in the queue dynamics. Balance between liquidity addition and removal suggests that we look at a heavy-traffic approximation for the HFT effects. We will first of all outline a heuristic picture, and then prove a rigorous result which is analogous to Lemma 3.1.

The essential scaling we wish to enforce can be seen from (1). Looking at (6), our model of LT-ers comes from the looking at the dominant (i.e., fluid) approximation of a Poisson process; i.e., (7). The next term in the expansion would be the diffusive (i.e., heavy-traffic) part of (6) given by the \( V_{L,H} \) terms. We want HFT traders, as opposed to the heavy traffic component of (6), to be the next term in the dynamics of the order book. Our motivation is that HFT-er’s should indeed play a perturbative but visible effect on the market. To reiterate some of our comments of Section 1, our goal is to build a stylized model where we can rigorously understand the transition from microstructure to volatility.

Fix a small parameter \( \varepsilon > 0 \). This will represent the time-scale at which HFT-ers operate. The defining characteristic of HFT-ers is that they react to the market at a much faster pace than the LT-ers. We will represent this time scale as \( \varepsilon \), and in a moment we will incorporate it into a model in a precise way. We will assume that HFT-ers add and remove liquidity in increments of \( \varepsilon \nu \), where \( \nu \) is a fixed positive exponent. Let’s next fix two stochastic processes \( \{ \lambda_t^+; t \geq 0 \} \) and \( \{ \lambda_t^-; t \geq 0 \} \) (which we will specify in a moment) and an exponent \( \gamma > 0 \). At time \( t \geq 0 \), the HFT-ers add liquidity with instantaneous rate \( e^{-\gamma \lambda_t^+} \) and remove liquidity at instantaneous rate \( e^{-\gamma \lambda_t^-} \). As with liquidity traders, we should think of \( \lambda_t^+ \) and \( \lambda_t^- \) as representing a statistical description of the behavior of the HFT community as a whole (and not the behavior of any single trader). We specify \( \lambda_t^+ \) and \( \lambda_t^- \) as follows. Fix \( \lambda, \sigma_+, \sigma_-, \alpha_+, \) and \( \alpha_- \) in \( (0, \infty) \). The dynamics of \( \lambda_t^+ \) are given by

\[
d\lambda_t^+ = -\frac{\alpha_+}{\varepsilon} (\lambda_t^+ - \bar{\lambda}) dt + \frac{\sigma_+}{\varepsilon} \sqrt{\lambda_t^+} dW_t^+, \quad t > 0
\]

\[
\lambda_0^+ = \lambda_0^+,
\]

where \( W^+ \) and \( W^- \) are two independent standard Brownian motions and \( \lambda_0^+ \) and \( \lambda_0^- \) in \( (0, \infty) \) are fixed. The dynamics of (14) are those of a CIR process, so standard theory as in Ikeda and Watanabe (1989) ensures that the SDE’s of (14) have unique nonnegative strong solutions. We also assume, for simplicity, that

\[
2\alpha_+ \bar{\lambda} > \sigma_+^2.
\]

This will simplify some of the estimates of Appendix D; physically it means that \( \lambda_t^+ \) are strictly positive and that thus the high frequency traders are always present in our model). We will develop some relevant bounds on CIR processes in Appendix D.

We will motivate these dynamics for the rates in a moment; at the moment, however, let’s complete our model of the order queue. We want both liquidity traders and high-frequency traders to be trading in the order queue. Remaining within the realm of Markovian dynamics for the order queue, we want to augment the generator of (3) by the high-frequency traders. For \( f \in C^\infty(\mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+) \), define

\[
(A_{\delta, \varepsilon} f)(x, \lambda_+, \lambda_-) = -\frac{\alpha_+}{\varepsilon} (\lambda_+ - \bar{\lambda}) \frac{\partial f}{\partial \lambda_+}
\]

\[
\times (x, \lambda_+, \lambda_-) + \frac{\sigma_+^2}{2\varepsilon} \lambda_+ \frac{\partial^2 f}{\partial \lambda_+^2} (x, \lambda_+, \lambda_-)
\]

\[
- \frac{\alpha_-}{\varepsilon} (\lambda_- - \bar{\lambda}) \frac{\partial f}{\partial \lambda_-} (x, \lambda_+, \lambda_-)
\]

\[
+ \frac{\sigma_-^2}{2\varepsilon} \lambda_- \frac{\partial^2 f}{\partial \lambda_-^2} (x, \lambda_+, \lambda_-)
\]

\[
+ \frac{1}{\varepsilon^\gamma} \left\{ f(x + \varepsilon \nu, \lambda_+, \lambda_-) - f(x, \lambda_+, \lambda_-) \right\} \lambda_+
\]

\[
+ \frac{1}{\varepsilon^\gamma} \left\{ f(x - \varepsilon \nu, \lambda_+, \lambda_-) - f(x, \lambda_+, \lambda_-) \right\} \lambda_-
\]

\[
+ \frac{1}{\delta_+^\gamma} \left\{ f(x + \delta_+ \nu, \lambda_+, \lambda_-) - f(x, \lambda_+, \lambda_-) \right\} \lambda_L^+
\]

\[
+ \frac{1}{\delta_-^\gamma} \left\{ f(x - \delta_- \nu, \lambda_+, \lambda_-) - f(x, \lambda_+, \lambda_-) \right\} \lambda_L^-
\]

(16)
Let’s also assume that the order queue has the same initial condition \( q^0 \) as in our LT model. Putting all of this together, let \( Q^\delta \) be a Markov process which starts at \( q^0 \) and which has generator given by (16); i.e., we suppose that

\[
f(\lambda^\delta, \lambda^+ \delta, \lambda^- \delta) = f(q^0, \lambda^0, \lambda^- ) + \int_{s=0}^{t} (A_\delta \cdot f)(Q^\delta, \lambda^+ \delta, \lambda^- \delta) ds + M_t
\]

where \( M \) is a martingale with respect to the filtration

\[
\mathcal{F}_t = \sigma(\{Q^\delta, \lambda^+ \delta, \lambda^- \delta; 0 \leq s \leq t\}, t \geq 0)
\]

As with \((4)\), we define \( \tau^\delta \) as

\[
\tau^\delta = \inf \{ t > 0 : Q^\delta_t = 0 \},
\]

and we assume that the order queue is given by \( \{Q^\delta, \lambda^\delta, \lambda^+ \delta, \lambda^- \delta; t \geq 0\} \).

Where does \((14)\) come from?

**Remark 4.1.** The processes \((14)\) are simply rescaled CIR processes. Let \( \lambda^\pm \) be the solutions of

\[
d\lambda^\pm_t = -\alpha_{\pm} (\lambda^\pm_t - \bar{\lambda}) dt + \sigma_{\pm} \sqrt{\lambda^\pm_t} dW^\pm_t + \alpha_{\pm} \lambda^\pm_t d\bar{\nu}_t,
\]

\[
\lambda^\pm_t = \lambda^\pm_0 + \int_{t=0}^{T_{\delta,\epsilon}} \sigma^\pm \sqrt{\lambda^\pm_s} dW^\pm_s + \alpha_{\pm} \lambda^\pm_s d\bar{\nu}_t,
\]

where \( \{\lambda^\pm_t; t \geq 0\} \) has the same law as \( \{\lambda^\pm; t \geq 0\} \). Thus \( \lambda^\pm \) is of order 1 and fluctuates with time scale of order \( \epsilon^2 \).

In other words, \( \lambda^+ \) and \( \lambda^- \) have relaxation time of order \( \epsilon \). We interpret this relaxation time as the time scale at which the HFT-ers react to the market.

**Remark 4.2.** A natural scaling relation is now apparent. We want the time scale of the HFT-ers to be much smaller than the time it takes for the midpoint to move (i.e., for the queue to empty). Namely, we want that

\[
\epsilon \ll \delta^2.
\]

In other words, we want the rates \( \lambda^\pm \) to reach equilibrium by time \( T_{\delta,\epsilon} \).

The import of using CIR processes is that they allow us a number of useful computations. To focus on the important issues, note that liquidity addition and removal due to the HFT-ers is given by

\[
\epsilon^{\nu} \hat{N}_t^\pm \left( \epsilon^{-\gamma} \int_{s=0}^{t} \lambda^\pm_t ds \right)
\]

where \( \hat{N}_t^+ \) and \( \hat{N}_t^- \) are two standard independent Poisson processes; i.e., liquidity addition and removal is governed by small Poisson processes run at clocks

\[
\epsilon^{-\gamma} \int_{s=0}^{t} \lambda^\pm_s ds.
\]

To understand these clocks, let’s rewrite \((14)\) as

\[
\left( \lambda^\pm - \bar{\lambda} \right) = \left( \lambda^\pm_0 - \bar{\lambda} \right) - \frac{\alpha_{\pm}}{\epsilon} \int_{s=0}^{t} (\lambda^\pm_s - \bar{\lambda}) ds + \frac{\sigma_{\pm}}{\sqrt{\epsilon}} \int_{s=0}^{t} \sqrt{\lambda^\pm_s} dW^\pm_s + \frac{\epsilon}{\alpha_{\pm}} (\lambda^\pm_0 - \lambda^\pm_0).
\]

We want to treat this as an expansion of \( \int_{s=0}^{t} \lambda^\pm_s ds \) in increasing powers of the small parameter \( \epsilon \). Let’s use (1) and (22) to approximate (20). We have that

\[
\epsilon^{\nu} \hat{N}_t^\pm \left( \epsilon^{-\gamma} \int_{s=0}^{t} \lambda^\pm_t ds \right)
\]

\[
\approx \epsilon^{\nu-\gamma} \int_{s=0}^{t} \lambda^\pm_0 ds + \epsilon^{\nu-\gamma} \sqrt{\lambda^\pm_0} \int_{s=0}^{t} \sqrt{\lambda^\pm_0} dW^\pm_s
\]

\[
+ \epsilon^{\nu-\gamma} \frac{\sigma_{\pm}}{\alpha_{\pm}} \int_{s=0}^{t} \sqrt{\lambda^\pm_s} dW^\pm_s
\]

\[
= \epsilon^{\nu-\gamma} \hat{N}_t^\pm \left( \epsilon^{-\gamma} \int_{s=0}^{t} \lambda^\pm_0 ds \right)
\]

\[
+ \epsilon^{\nu-\gamma} \frac{\sigma_{\pm}}{\alpha_{\pm}} \int_{s=0}^{t} \sqrt{\lambda^\pm_0} dW^\pm_s
\]

\[
+ \epsilon^{\nu-\gamma} \hat{N}_t^\pm \left( \epsilon^{-\gamma} \int_{s=0}^{t} \lambda^\pm_0 ds \right)
\]

\[
+ \epsilon^{\nu-\gamma} \frac{\sigma_{\pm}}{\alpha_{\pm}} \int_{s=0}^{t} \sqrt{\lambda^\pm_0} dW^\pm_s
\]

where \( \hat{V}_a^\pm \) and \( \hat{V}_b^\pm \) are some other Brownian motions.
The third line of (23) comes from rewriting the $dW^\pm$ integrals (which are martingales) as time-changed Brownian motions (Karatzas & Shreve (1991, Theorem 3.4.6) provides a review of this), and the final line comes from using the first term of (22).

We can now quantify the behavior of the HFT effects. Define

$$
\varsigma \equiv \min \left\{ \frac{1}{2} + \nu - \gamma, \nu - \gamma/2 \right\}.
$$

(24)

This will be the macroscopically visible scale of the HFT behavior (as the dominant correction to the fluid approximation of the LT-ers); asymptotically

$$
\max \{ \varepsilon^{\nu-\gamma+1/2}, \varepsilon^{\nu-\gamma/2} \} \ll \varepsilon^{\varsigma}
$$

as $\varepsilon \searrow 0$. There are three regimes:

- **Regime 1:** $\gamma < 1$. Here $\varsigma = \nu - \gamma/2$ so $\varepsilon^{\nu-\gamma+1/2} \ll \varepsilon^{\nu-\gamma/2}$. Here the dominant stochasticity comes from the $V^a,\pm,\varepsilon$ terms; i.e., from the Poisson stochasticity of the arrivals of orders, cancellations, and executions.

- **Regime 2:** $\gamma > 1$. Here $\varsigma = 1 + \nu - \gamma$ so $\varepsilon^{\nu-\gamma+1/2} \ll \varepsilon^{\nu-\gamma/2}$. Here the dominant stochasticity comes from the $V^b,\pm,\varepsilon$ terms; i.e., from the stochasticity of the rates at which liquidity is added and removed.

- **Regime 3:** $\gamma = 1$. Here $\varsigma = 1/2 + \nu - \gamma = \nu - \gamma/2$ so $\varepsilon^{\nu-\gamma+1/2} = \varepsilon^{\nu-\gamma/2}$. Here the $V^a,\pm,\varepsilon$ and $V^b,\pm,\varepsilon$ sources of stochasticity are present at the same level.

Let’s also define

$$
\varsigma^2 \equiv \begin{cases} 
2 \left( \frac{\sigma_a}{\alpha_v} \right)^2 + \left( \frac{\sigma_b}{\alpha_v} \right)^2 & \text{if } \gamma < 1 \\
\left( \frac{\sigma_a}{\alpha_v} \right)^2 + \left( \frac{\sigma_b}{\alpha_v} \right)^2 & \text{if } \gamma = 1 \\
\left( \frac{\sigma_a}{\alpha_v} \right)^2 & \text{if } \gamma > 1.
\end{cases}
$$

We set

$$
\chi^{\pm,\varepsilon} \equiv \frac{\varepsilon^{2\nu-\gamma + \varepsilon + 2(\nu - \gamma)} \left( \frac{\sigma_a}{\alpha_v} \right)^2}{\varepsilon^\varsigma \sigma_{\chi}}.
$$

and

$$
\chi^{\varepsilon} \equiv \sqrt{\chi^{+,\varepsilon} + \chi^{-,\varepsilon}}
$$

(25)

Note that $\lim_{\varepsilon \searrow 0} \chi_{\varepsilon} = 1$ and that $\tilde{\chi}$ is finite. Then the asymptotics of (23) can be written as

$$
\varepsilon^{\nu} \tilde{N}^\pm \left( \varepsilon^{-\gamma} \int_{s=0}^t \lambda^{\pm,\varepsilon}_s ds \right) \approx \varepsilon^{\nu-\gamma} \tilde{\lambda} t
$$

$$
+ \varepsilon^\varsigma \tilde{\chi}^{\pm,\varepsilon} \sqrt{\lambda} \tilde{V}^{+,\varepsilon}
$$

where $\tilde{V}^{+,\varepsilon}$ is a Brownian motion. The combined effect of liquidity addition and removal will thus be

$$
\varepsilon^{\nu} \tilde{N}^+ \left( \varepsilon^{-\gamma} \int_{s=0}^t \lambda^{+\varepsilon}_s ds \right)
$$

$$
- \varepsilon^{\nu} \tilde{N}^- \left( \varepsilon^{-\gamma} \int_{s=0}^t \lambda^{-\varepsilon}_s ds \right)
$$

$$
\approx \varepsilon^\varsigma \tilde{\chi}^{+\varepsilon} \sqrt{\lambda} \tilde{V}^{+,\varepsilon} - \varepsilon^\varsigma \tilde{\chi}^{-\varepsilon} \sqrt{\lambda} \tilde{V}^{-,\varepsilon}
$$

$$
= \varepsilon^\varsigma \tilde{\chi}^{\varepsilon} \sqrt{\lambda} \tilde{V}^{d,\varepsilon}
$$

(27)

where $\tilde{V}^{d,\varepsilon}$ is yet another standard Brownian motion.

We can now start to combine things together. Adding the HFT approximation of (27) to the fluid approximation (7) of the liquidity traders, we get that

$$
Q_t^{\varepsilon} \approx q^{\hat{t}} - \delta^{\varepsilon} - \gamma + \varepsilon^{\nu-\gamma} \hat{t} + \varepsilon^\varsigma \tilde{\chi}^{d,\varepsilon}.
$$

(28)

We now have a problem which we can connect to well-known theory; to find when the queue empties (and thus when the midpoint price changes), we should understand when a Brownian motion with drift hits a point. Since the Brownian motion is small, we have an extra complication that the analysis should be asymptotic; nevertheless, we do have a starting point.

**Remark 4.3.** In much the same way that stochastic volatility has opened the door to broad collection of useful corrections to the basic model of asset prices as geometric Brownian motion, the stochasticity of $\lambda^{h,\varepsilon}$ provides extra degrees of freedom for modelling high-frequency trading. In particular, our stochastic rate of order placement allows for periods of very high order placement (one might think of the Flash Crash as an example). The two regimes of $\gamma < 1$ and $\gamma > 1$ correspond to two different situations. In the first (where $\gamma < 1$), the dominant effect on volatility is...
due to the randomness between times at which HFT-ers add or remove liquidity. In the second (where \( \gamma > 1 \)),
the dominant effect is due to the stochasticity of the order rate; i.e., to periods during which liquidity is added or removed
at abnormally high rates.

Remark 4.4. As we mentioned near the beginning of the section, we want (28) to reflect the dominant two terms in the
description of the order queue dynamics. Comparing (28) and (6), this means that we want

\[ \varepsilon^2 \gtrsim \delta^{\nu_{\gamma} - \gamma}/2. \]  

(29)

Remark 4.5. To see the essential issues, define

\[ X_t = x - bt + \varepsilon W_t \]  

(30)

where \( W \) is a Brownian motion and \( \varepsilon \in (0, \infty) \) is a fixed constant. We are interested in the asymptotics of

\[ \tau^* = \inf \{ t > 0 : X_t \leq 0 \}. \]

In particular, if \( \varepsilon = 0 \), then \( \tau^* = \tau^*_0 \). We are interested in \( \tau^* - \tau^*_0 \) for small \( \varepsilon \). Fix \( \alpha_* > 0 \) and define

\[ \xi^{\alpha_*} = \frac{\tau^* - \tau^*_0}{\alpha_*}. \]

Fix \( \theta \in \mathbb{R} \) and consider the process

\[ Z_t = \exp \left[ AX_t + i\theta \left( \frac{t - \tau^*}{\alpha_*} \right) \right] \quad t \geq 0 \]

where \( i = \sqrt{-1} \). Since \( X_t \geq 0 \) for \( t \leq \tau^* \), we can take the expectation of (30) and use optional sampling
to see that

\[ \mathbb{E} [\tau^* \wedge t] \leq \frac{x}{b} \]

for all \( t \geq 0 \), so \( \mathbb{E} [\tau^*] = \lim_{\tau^* \rightarrow \infty} \mathbb{E} [\tau^* \wedge t] \leq \frac{x}{b} < \infty \) so \( \mathbb{P} \{ \tau^* < \infty \} = 1 \). Assuming that \( Z \) is a martingale, we have that

\[ \mathbb{E} [\exp \{ i\theta \xi^{\alpha_*} \}] = \mathbb{E} \left[ \exp \left[ i\theta \left( \frac{\tau^* - \tau^*_0}{\alpha_*} \right) \right] \right] = \exp \left[ AX - \frac{i\theta \tau^*_0}{\alpha_*} \right]. \]

But for \( Z \) to be a martingale, we must have that

\[ \frac{\varepsilon^2}{2} A^2 - bA + \frac{i\theta}{\alpha_*} = 0. \]  

(31)

In other words, we must have that

\[ A = \frac{b - \sqrt{b^2 - 4 \left( \frac{\varepsilon^2}{2} \right) \left( \frac{i\theta}{\alpha_*} \right)}}{2 \left( \frac{\varepsilon^2}{2} \right)} = \frac{b - \sqrt{b^2 - 2ix^2 \theta}}{\varepsilon^2} \]

(taking here the principal branch of the square root function on \( \mathbb{C} \)). We take here the negative square root to reflect the requirement that \( A \) vanish when \( \theta \) vanishes. Using three terms in the asymptotic expansion of \( f(x) \approx \sqrt{1 - x} \), we have that

\[ \frac{x^2}{b \alpha_*} \ll 1 \]  

(32)

then

\[ A \approx \frac{b}{\varepsilon^2} \left\{ 1 - \left( 1 - \frac{1}{2} \left( \frac{2ix^2 \theta}{b^2 \alpha_*} \right) - \frac{1}{8} \left( \frac{2ix^2 \theta}{b^2 \alpha_*} \right)^2 \right) \right\} \]

\[ = \frac{b}{\varepsilon^2} \left\{ \frac{1}{2} \left( \frac{2ix^2 \theta}{b^2 \alpha_*} \right) + \frac{1}{8} \left( \frac{2ix^2 \theta}{b^2 \alpha_*} \right)^2 \right\} \]

\[ = \frac{it \theta}{b \alpha_*} - \frac{x^2}{2b \alpha_*^2} \theta^2. \]  

(33)

Thus if (32) holds

\[ Ax - \frac{i\theta x}{\alpha_*} \approx \frac{x^2}{b \alpha_*} - \frac{x^2}{2b \alpha_*} \theta^2 - \frac{x^2}{b \alpha_*} = -\frac{x^2}{2b \alpha_*} \theta^2. \]

We are interested in the case where \( x \) is in some sense small, so in order for \( \xi^{\alpha_*} \) to have a nontrivial limit as
\( x, b, \) and \( \varepsilon \) fluctuate, we want to select

\[ \alpha_* \overset{\text{def}}{=} \frac{x}{b} \sqrt{\varepsilon}. \]  

(34)

Putting this back in (32), we want that

\[ \frac{x^2}{b^3 \varepsilon^{3/2}} \ll 1 \]

or rather that

\[ \frac{x}{\sqrt{bx}} \ll 1. \]  

(35)
Under these assumptions, \( \xi^\alpha_{\epsilon} \) is approximately Gaussian with zero mean and unit variance.

To capture the relevant scales in our case, we want to take

\[
x = \delta^{\alpha}, \quad b = \delta^{\nu - \gamma}, \quad \text{and} \quad \varepsilon = \varepsilon \delta
\]  

(since we are interested in scales, and not pre factors, we here set \( q = \ell = 1 \)).

The analogue of \( \alpha \) is thus

\[
\varepsilon^{\delta} \sqrt{\delta^n} = \delta^{1 - \nu + \gamma} \varepsilon^{\delta}
\]

The analogue of (35) is that

\[
\varepsilon^{\delta} \ll \sqrt{\delta^n} \delta^{\nu - \gamma} = \delta^{1 - \nu + \gamma} \varepsilon^{\delta}
\]  

Let’s now organize our asymptotic relations between \( \delta \) and \( \varepsilon \). Recall that our only stipulation at the moment on the various exponents is (13) (and of course we defined \( \iota \) via (10)). The relation between \( \delta \) and \( \varepsilon \) is at the moment given by (19), (29), and (37). To these we will add a technical requirement that

\[
\frac{\varepsilon^{\delta}}{\delta^{1 - \nu + \gamma}} \ll 1.
\]  

To organize these asymptotic requirements in a clear way, let’s use the notion of directed systems (Royden 1988, Section 8.7) is a good introduction to these concepts.

**Definition 4.6.** We say that \( D(\nu, \gamma) \) is the collection of subsets \( A \) of \( \mathbb{R}^2 \) of the form

\[
\Theta \cap (0, 1)^2
\]

where \( \Theta \) is a neighborhood of the origin \((0, 0)\) and where

\[
\sup_{(\delta, \varepsilon) \in A} \frac{\varepsilon^{(\gamma + 1)/\delta}}{\delta} < \infty,
\]

\[
\inf_{(\delta, \varepsilon) \in A} \frac{\varepsilon^{\delta}}{\delta^{1 + \nu - \gamma}} > 0,
\]

and

\[
\sup_{(\delta, \varepsilon) \in A} \frac{\varepsilon^{\delta}}{\delta^{1 + \nu - \gamma}} < \infty.
\]  

If \( A_1 \) and \( A_2 \) are in \( D(\nu, \gamma) \), we say that \( A_2 \geq_{(\gamma, \nu)} A_1 \) if \( A_2 \subset A_1 \).

Recall that if \( X(\delta, \varepsilon) \) takes values in a topological space \( S \) for all \( \delta \) and \( \varepsilon \) in \((0, 1)\), then

\[
\lim_{(\delta, \varepsilon) \in A} X(\delta, \varepsilon) = \bar{X}
\]

for some \( \bar{X} \in S \) if for every neighborhood \( O \) of \( \bar{X} \) there is an \( A_0 \in D(\nu, \gamma) \) such that \( X(\delta, \varepsilon) \in O \) for all \((\delta, \varepsilon) \in A \in D(\nu, \gamma) \) where \( A \geq A_0 \). For example, we have that

\[
\lim_{(\delta, \varepsilon) \in A} \frac{\varepsilon^{(\gamma + 1)/\delta}}{\delta} = 0,
\]

\[
\lim_{(\delta, \varepsilon) \in A} \frac{\varepsilon^{\delta}}{\delta^{1 + \nu - \gamma}} = \infty,
\]

and

\[
\lim_{(\delta, \varepsilon) \in A} \frac{\varepsilon^{\delta}}{\delta^{1 + \nu - \gamma}} = 0.
\]  

**Remark 4.7.** We have directly copied (29) and (37) into (39), but where are (19) and (38)? It turns out that

\[
\varepsilon^{(\gamma + 1)/\delta} \ll \delta
\]  

(which corresponds to the first requirement of (40)) if and only if (19) and (38) hold. Indeed, (19) and (38) hold if and only if

\[
\max\{\varepsilon^{\gamma}, \varepsilon^{\nu - \gamma}\} \ll \delta
\]

which holds if and only if

\[
\varepsilon^{\min\{1/2, \nu - \gamma\}} \ll \delta.
\]  

But

\[
\min\{1/2, \nu - \gamma\} = \begin{cases} \min\{1/2, \nu - (\nu - \gamma)/2\} & \text{if } \gamma \leq 1 \\ \min\{1/2, \nu - (1/2 + \nu - \gamma)\} & \text{if } \gamma > 1 \end{cases}
\]

\[
\min\{1/2, \gamma/2\} = \begin{cases} \min\{1/2, \gamma - 1/2\} & \text{if } \gamma \leq 1 \\ 1/2 & \text{if } \gamma > 1 \end{cases} = (\gamma \land 1)/2.
\]

Thus (41) indeed holds if and only if (19) and (38) hold.

5. **Limit theorem for price change**

We want to here rigorously extract asymptotics of the law of \( \tau^{\delta, \varepsilon} \) from (17). We will use a variant of Lemma 3.1 in order to rigorously achieve the desired
results under minimal assumptions. Define
\begin{equation}
\xi_{\delta,\varepsilon} \overset{\text{def}}{=} \frac{\tau_{\delta,\varepsilon} - \hat{T}\delta^2}{\delta^{1-\nu_\varepsilon + \gamma_\varepsilon} \varepsilon^\varepsilon}.
\end{equation}

Then
\begin{equation}
\tau_{\delta,\varepsilon} = \delta^2 \hat{T} + \delta^{1-\nu_\varepsilon + \gamma_\varepsilon} \varepsilon^\varepsilon \xi_{\delta,\varepsilon} = \delta^2 \hat{T} \Theta(\delta, \varepsilon, \xi_{\delta,\varepsilon})
\end{equation}

where we have for convenience defined
\begin{equation}
\Theta(\delta, \varepsilon, R) \overset{\text{def}}{=} 1 + \left( \frac{\varepsilon^\varepsilon}{\delta^{1+\nu_\varepsilon - \gamma_\varepsilon}} \right) \frac{R}{\hat{T}}
\end{equation}

for all \( \delta \) and \( \varepsilon \) in \((0, 1)\) and all \( R \in \mathbb{R} \). Then (37) (which stems form the third requirement of (39)) implies that \( \xi_{\delta,\varepsilon} \) is a small component of \( \tau_{\delta,\varepsilon} \). Our main result is an asymptotic description of the law of \( \xi_{\delta,\varepsilon} \).

**Proposition 5.1.** We have that
\begin{equation}
\lim_{(\delta,\varepsilon) \in A} \xi_{\delta,\varepsilon} = \frac{\sigma \sqrt{\lambda q}}{\ell^{3/2}} G = \frac{\sigma \sqrt{\lambda T}}{\ell} G,
\end{equation}

where \( G \) is a standard Gaussian random variable, this being understood in the sense of weak convergence.

6. Simulation

Simulations verify the claimed asymptotics of our model. Figure 1 is a typical simulation of the system (the combination of liquidity traders and high-frequency traders). To model the liquidity traders, we take \( \delta = .3, \nu_\varepsilon = 0.001 \) and \( \gamma_\varepsilon = 5 \). Then \( \ell = 0.001 - 5 + 2 = -2.999 \). Let’s take \( q_\varepsilon = 100, \lambda^+_\varepsilon = 1 \) and \( \lambda^-_\varepsilon = 9 \). Then the initial size of the order queue is \( q_\varepsilon \delta^2 = 3699 \). Liquidity trades add to the queue at rate \( \lambda^+_\varepsilon \delta^{\nu_\varepsilon - \gamma_\varepsilon} = 1.82 \) and removing (through execution or removal of orders) at rate \( \lambda^-_\varepsilon \delta^{\nu_\varepsilon - \gamma_\varepsilon} = 16.4 \) (i.e., \( \lambda^+_\varepsilon = 1 \) and \( \lambda^-_\varepsilon = 9 \), \( \nu_\varepsilon = .001 \) and \( \gamma_\varepsilon = 5 \)).

Calculating that
\begin{equation}
\hat{T} = \frac{q_\varepsilon}{\lambda^-_\varepsilon - \lambda^+_\varepsilon} = \frac{100}{9 - 1} = 12.5,
\end{equation}

the order queue should thus empty at time
\begin{equation}
\hat{T} \delta^2 = 12.5(3)^2 = 1.125.
\end{equation}

The stochasticity of the liquidity and high frequency traders is then overlaid on top of this expected time to empty.

We can also numerically verify Proposition 5.1. Figure 2 contains histograms of \( \xi_{\delta,\varepsilon} \) based on 1000 samples of realizations as in Figure 1. We have used most of the parameters of Figure 1. We have used either \( \lambda = 40 \) or \( \lambda = 100 \). We have also taken either \( \lambda^+_\varepsilon = 9 \) (which gives (44)) or \( \lambda^-_\varepsilon = 2 \) (which gives \( T = 100 \)).

The red curve in each histogram is an approximate normal distribution line based on the data, and the shape of the histogram of 1000 sample data approximates the shape of the red curve, verifying Proposition 5.1, that \( \xi_{\delta,\varepsilon} \) is approximately normal. The green curves are the corresponding Gaussians with the theoretical variance as given by Proposition 5.1.

---

**Fig. 1.** Simulation result of Order Queue Size. Parameters are \( \delta = .3, \varepsilon = .01, \lambda^+_\varepsilon = 1, \lambda^-_\varepsilon = 9, \lambda = 40, \nu_\varepsilon = .001, \gamma_\varepsilon = 5, \alpha_1 = .065, \alpha_2 = .061, \lambda_0 = 41, \sigma_1 = .057, \sigma_2 = .0576, \nu = .001, \) and \( \gamma = 1.5 \). Dotted line is theoretical average behavior of queue.
Fig. 2. Histograms of Simulations. Common parameters are $\delta = 0.3$, $\epsilon = 0.01$, $\lambda^+ = 1$, $\nu_0 = 0.001$, $\gamma_0 = 5$, $\alpha_1 = 0.065$, $\alpha_2 = 0.061$, $\lambda_0 = \lambda + 1$, $\sigma_1 = 0.057$, $\sigma_2 = 0.0576$, $\nu = 0.001$, and $\gamma = 1.5$.

Table 1

<table>
<thead>
<tr>
<th>Times</th>
<th>$\lambda = 40$, $T = 12.5$</th>
<th>$\lambda = 40$, $T = 100$</th>
<th>$\lambda = 100$, $T = 12.5$</th>
<th>$\lambda = 100$, $T = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical Mean</td>
<td>0.252</td>
<td>2.095</td>
<td>1.059</td>
<td>7.407</td>
</tr>
<tr>
<td>Theoretical Variance</td>
<td>13.043</td>
<td>6677.842</td>
<td>32.607</td>
<td>16694.61</td>
</tr>
<tr>
<td>Empirical Variance</td>
<td>12.641</td>
<td>6725.253</td>
<td>36.629</td>
<td>16439.18</td>
</tr>
<tr>
<td>Relative Error</td>
<td>3.08%</td>
<td>0.71%</td>
<td>12.34%</td>
<td>-1.53%</td>
</tr>
</tbody>
</table>

Fig. 3. Poisson arrival rate for HFT; parameters are $\lambda_0 = \bar{\lambda} = 40$, $\lambda_L = 9$, $\lambda_L^+ = 1$, $\delta = 0.3$, $\epsilon = 0.01$, $\nu_0 = 0.001$, $\gamma_0 = 5$, $\alpha_1 = \alpha_2 = 0$, and $\sigma_1 = \sigma_2 = 0$, $\nu = 0.001$, and $\gamma = 1.5$.

The empirical variance matches that of the theoretical variance with small relative error. Table 1 gives simulated and theoretical variance of the times at which the queues empty.

We note that the bottom two histograms of Figure 2 (where $\lambda^-_L = 2$) are noticeably asymmetric, more so than the top two histograms (where $\lambda^-_L = 9$). The smaller the value of $\lambda^-_L$, the longer it takes for the queue to empty due to the liquidity traders (i.e., $\hat{T}$ is larger when $\lambda^-_L$ is closer to $\lambda^+_L$). The longer this time is, the more time that the high frequency traders have to reach a (dynamic) equilibrium. More precisely, the larger $\hat{T}$ is, the more the arrival rate of HFT orders looks like a gamma distribution (the invariant measure of a CIR process). In Figure 3, we set $\lambda_0 = \lambda$ and $\sigma_+ = \sigma_- = 0$, so that the HFT orders are purely Poisson. The distribution of $\xi_{\lambda,\epsilon}$ in this case is much more symmetric.
7. From epochs to volatility

7.1. Asymptotics of queues

What have we done thus far? We have quantified the asymptotics of the times at which a typical order queue empties. Where do we go from here? Let’s now separately consider the bid and ask queues. For reference, let’s assume that the midpoint price has changed at time 0−, so our model restarts at time 0.

We want to model an unbiased market; i.e., where liquidity traders add liquidity to both queues in \( \delta \) and that liquidity traders add liquidity to both queues in increments of \( \delta \). Assume that \( \delta \) is small enough that the bid queue is governed by

\[
\hat{\gamma} \left( \delta \right) = \min \left( \hat{\gamma} \left( \delta \right) - \delta^2 \right) \left( \delta \right),
\]

the time at which the midpoint price will again change. If \( \tau^A,\delta,\epsilon < \tau^B,\delta,\epsilon \), then the midpoint price moves down, and if \( \tau^B,\delta,\epsilon > \tau^A,\delta,\epsilon \), the midpoint price moves up. Clearly \( \lim_{\delta,\epsilon \to 0} \hat{\gamma}(\delta, \epsilon) = 0 \). Setting

\[
\hat{\mu}_{\delta,\epsilon}(S) \overset{\text{def}}{=} \mathbb{P}(\hat{\xi}_{\delta,\epsilon} \in S); \quad S \in \mathcal{B}(\mathbb{R})
\]

we want to consider an asymptotic regime where \( \hat{\mu}_{\delta,\epsilon} \) has a limit. Our goal is to then reconstruct the volatility of the asset price from the limit of \( \hat{\mu}_{\delta,\epsilon} \).

Let’s now construct another directed system. Define the one-dimensional simplex

\[
P^* = \left\{ (p_B, p_A) \in [0,1]^2 : p_B + p_A = 1 \right\}.
\]

Fix \( P^* = (p_B, p_A) \in P \).

**Definition 7.1.** Let \( \mathcal{D}(P^*) \) be the collection of sets \( C \in (0,1)^2 \) such that \( C \in \mathcal{D}(\nu_B, \gamma_B) \cap \mathcal{D}(\nu_A, \gamma_A) \) and let

\[
\mathcal{D} \cap P(46)
\]

we have that in an appropriate asymptotic regime (given in Proposition 5.1), \( \xi_{\delta,\epsilon} \) has a well-defined weak limit. Assuming that liquidity behavior for the bid queue is independent of that of the ask queue, \( \xi_{B,\delta,\epsilon} \) and \( \xi_{A,\delta,\epsilon} \) will be independent. To proceed, let’s define

\[
\hat{\gamma}(\delta, \epsilon) \overset{\text{def}}{=} \hat{\gamma}(\delta, \epsilon) + \gamma(\delta, \epsilon)
\]

\[
\hat{\xi}_{\delta,\epsilon} \overset{\text{def}}{=} \min \left\{ \hat{\gamma}(\delta, \epsilon) \xi_{B,\delta,\epsilon}, \hat{\gamma}(\delta, \epsilon) \xi_{A,\delta,\epsilon} \right\}
\]

where

\[
\hat{\xi}_{\delta,\epsilon} \overset{\text{def}}{=} \xi_{\delta,\epsilon} - \gamma(\delta, \epsilon)
\]

\[
\hat{\gamma}(\delta, \epsilon) \overset{\text{def}}{=} \gamma(\delta, \epsilon) + \gamma(\delta, \epsilon)
\]

\[
\hat{\xi}_{\delta,\epsilon} \overset{\text{def}}{=} \min \left\{ \hat{\gamma}(\delta, \epsilon) \xi_{B,\delta,\epsilon}, \hat{\gamma}(\delta, \epsilon) \xi_{A,\delta,\epsilon} \right\};
\]

\( \tau^\delta_{\epsilon,\delta} \) is the time at which the midpoint price will again change. If \( \tau^\delta_{\epsilon,\delta} < \tau^\delta_{\epsilon,\delta} \), then the midpoint price moves down, and if \( \tau^\delta_{\epsilon,\delta} > \tau^\delta_{\epsilon,\delta} \), the midpoint price moves up. Clearly \( \lim_{\delta,\epsilon \to 0} \hat{\gamma}(\delta, \epsilon) = 0 \). Setting

\[
\hat{\mu}_{\delta,\epsilon}(S) \overset{\text{def}}{=} \mathbb{P}(\hat{\xi}_{\delta,\epsilon} \in S); \quad S \in \mathcal{B}(\mathbb{R})
\]

we want to consider an asymptotic regime where \( \hat{\mu}_{\delta,\epsilon} \) has a limit. Our goal is to then reconstruct the volatility of the asset price from the limit of \( \hat{\mu}_{\delta,\epsilon} \).

Let’s now construct another directed system. Define the one-dimensional simplex

\[
P^* = \left\{ (p_B, p_A) \in [0,1]^2 : p_B + p_A = 1 \right\}.
\]

Fix \( P^* = (p_B, p_A) \in P \).

**Definition 7.1.** Let \( \mathcal{D}(P^*) \) be the collection of sets \( C \in (0,1)^2 \) such that \( C \in \mathcal{D}(\nu_B, \gamma_B) \cap \mathcal{D}(\nu_A, \gamma_A) \) such that

\[
\left\{ \left( \frac{\gamma_{B}(\delta, \epsilon)}{\hat{\gamma}(\delta, \epsilon)}, \frac{\gamma_{A}(\delta, \epsilon)}{\hat{\gamma}(\delta, \epsilon)} \right) : (\delta, \epsilon) \in C \right\} = \mathcal{O} \cap P
\]
for some open subset \( \mathcal{O} \) of \( \mathbb{R}^2 \) (i.e., the set on the left in (46) is an open subset of \( \mathcal{P} \) in the topology which \( \mathcal{P} \) inherits from \( \mathbb{R}^2 \)), and such that
\[
\sup_{\delta, \varepsilon \in A} \frac{\delta^4}{\mathcal{H}(\delta, \varepsilon)} < \infty.
\]

If \( C_1 \) and \( C_2 \) are in \( \hat{\mathcal{D}}(P^*) \), we say that \( C_2 \geq C_1 \) if \( C_2 \geq_{(\nu_A, \gamma_A)} C_1 \) and \( C_2 \geq_{(\nu_B, \gamma_B)} C_1 \) (i.e., if \( C_2 \subset C_1 \)).

With this definition, we have that
\[
C \leq \mathcal{D}(P^*) \quad \text{and} \quad C \geq \mathcal{D}(P^*) \quad \text{from (47).}
\]

Note also that since
\[
\mathcal{H}(\delta, \varepsilon) < \infty
\]

for \( s \in \{A, B\} \) (using the definition of \( \mathcal{D}(\nu_s, \gamma_s) \)),
\[
\lim_{\delta, \varepsilon \in A} \frac{\mathcal{H}(\delta, \varepsilon)}{\delta^2} = 0.
\]

Comparing this with the second requirement of (47), we want that
\[
\delta^4 \ll \mathcal{H}(\delta, \varepsilon) \ll \delta^2.
\]

Let’s formalize some consequences of weak convergence under \( \mathcal{D}(P^*) \). Let \( \mathcal{Y}_B \) and \( \mathcal{G}_A \) be two independent standard Gaussian random variables. The first is that we asymptotically have equal likelihoods of increases and decreases of the midpoint price.

**Corollary 7.2.** We have that
\[
\lim_{\delta, \varepsilon \in A} \mathbb{P}\{\mathcal{Y}_B = \tau_{B}^{\delta, \varepsilon}\} = \mathbb{P}\{\mathcal{Y}_A = \tau_{A}^{\delta, \varepsilon}\} = \frac{1}{2}.
\]

The proof is given in Appendix C. The second is an explicit characterization of the limit of the law of \( \xi_{\delta, \varepsilon} \).

Define
\[
\hat{\mu}(S) \overset{\text{def}}{=} \mathbb{P}\{\min\{p_B V_B G_A, p_A V_A G_A\} \in S\}.
\]

**Corollary 7.3.** We have that
\[
\lim_{\delta, \varepsilon \in A} \hat{\mu}_{\delta, \varepsilon} = \hat{\mu}.
\]

Again, the proof is delayed until Appendix C.

Let’s now return to a macroscopic scale and consider a sequence of midpoint price changes like in (45). Namely, let \( \{\mathcal{L}_{n}^{\delta, \varepsilon}\}_{n \in \mathbb{N}} \) be an i.i.d. sequence of random variables with common law \( \hat{\mu}_{\delta, \varepsilon} \). Let \( \{\tau_{n}^{\delta, \varepsilon}\}_{n \in \mathbb{N}} \) be the collection of times defined as
\[
\tau_{0}^{\delta, \varepsilon} = 0
\]
\[
\tau_{n+1}^{\delta, \varepsilon} = \tau_{n}^{\delta, \varepsilon} + \mathcal{L}_{n+1}^{\delta, \varepsilon}.
\]

Let now \( \{q_n\}_{n \in \mathbb{N}} \) be an i.i.d. collection of random variables with
\[
\mathbb{P}\{q_n = 1\} = \mathbb{P}\{q_n = -1\} = \frac{1}{2}.
\]

We want to recursively define a martingale by having jumps of size \( \delta q_n \) at time \( \tau_{n+1}^{\delta, \varepsilon} \). Set
\[
M_{0}^{\delta, \varepsilon} = 0
\]
for \( 0 \leq t < \tau_{1}^{\delta, \varepsilon} \) and recursively define
\[
M_{t}^{\delta, \varepsilon} = M_{t_{n}}^{\delta, \varepsilon} + \delta q_n
\]
for all \( t \in [\tau_{n}, \tau_{n+1}) \). Then \( M^{\delta,\varepsilon} \) is right-continuous and has left-hand limits, and is a martingale. We want to think of \( M^{\delta,\varepsilon} \) as the martingale part of the logarithmic price of an asset. Note that our scaling will indeed give diffusive dynamics.

Important information about martingales can be gleaned from their quadratic variation. For each \( t \geq 0 \), define

\[
\left\langle M^{\delta,\varepsilon} \right\rangle_t \overset{\text{def}}{=} \lim_{\max_{1 \leq j \leq l} |t_j - t_{j-1}| \to 0} \sum_{j=1}^{l} \left( M_{t_j}^{\delta,\varepsilon} - M_{t_{j-1}}^{\delta,\varepsilon} \right)^2 
\]

(note that since \( M \) is discontinuous, \( M^{\delta,\varepsilon} \) cannot be understood as a Brownian motion run at the “clock” \( \left\langle M^{\delta,\varepsilon} \right\rangle \); nevertheless, we will use \( \left\langle M^{\delta,\varepsilon} \right\rangle \) to microscopically define the volatility of the price). We have that

\[
\left\langle M^{\delta,\varepsilon} \right\rangle_t = n \delta^2
\]

if \( t \in [\tau_{n}, \tau_{n+1}) \). We want to understand the asymptotic behavior of \( \left\langle M^{\delta,\varepsilon} \right\rangle \) for \( (\delta, \varepsilon) \in A \in \mathcal{D}(P^*) \).

Since the average time between trades is of order \( \delta^2 T \), \( \left\langle M^{\delta,\varepsilon} \right\rangle \) increases by \( \delta^2 \) in time \( \delta^2 T \), so we expect that

\[
\left\langle M^{\delta,\varepsilon} \right\rangle_t \approx \frac{t}{T}.
\]

Our interest is to see how high-frequency trading affects this asymptotic.

To proceed, let’s establish some control over the first two moments of \( \xi^{\delta,\varepsilon} \).

**Lemma 7.4.** For every \( A \in \mathcal{D}(P^*) \), we have that

\[
K_{A,7.4} \overset{\text{def}}{=} \sup_{(\delta, \varepsilon) \in A} \int_{x \in \mathbb{R}} |x|^2 \mu^{\delta,\varepsilon}(dx) < \infty.
\]

**Proof.** For any \( \lambda \in [0, 1] \) and \( x \) and \( y \) in \( \mathbb{R} \),

\[
| (\lambda x) \wedge ((1 - \lambda) y) |^2 \leq x^2 + y^2.
\]

Use Lemma B.1. \( \square \)

Thus for each \( \delta \) and \( \varepsilon \) in \( (0, 1) \),

\[
m^{\delta,\varepsilon} \overset{\text{def}}{=} \int_{x \in \mathbb{R}} x \mu^{\delta,\varepsilon}(dx)
\]

is well-defined. Let’s also define

\[
m \overset{\text{def}}{=} \int_{x \in \mathbb{R}} \bar{x} \mu(dx);
\]

since Gaussians have first moments, \( m \) is clearly well-defined. We then can identify the asymptotic of \( m^{\delta,\varepsilon} \).

**Lemma 7.5.** We have that

\[
\lim_{\varepsilon \to 0} \sup_{\delta \geq A} \frac{1}{\varepsilon} \left\{ \left\langle M^{\delta,\varepsilon} \right\rangle_t - \frac{t}{T} \right\} = Z^\delta = \frac{m}{\varepsilon} \left\{ \langle M^{\delta,\varepsilon} \rangle_t - \frac{t}{T} \right\}.
\]

We can now state our main result, which is a quantification of the correction of (49) due to the HFT- ers. Define

\[
Z^\delta_t = \frac{1}{2} \varepsilon (\delta, \varepsilon) \varepsilon \left\{ \langle M^{\delta,\varepsilon} \rangle_t - \frac{t}{T} \right\},
\]

\[
\hat{Z}_t \overset{\text{def}}{=} \frac{m}{\varepsilon} \left\{ \langle M^{\delta,\varepsilon} \rangle_t - \frac{t}{T} \right\}.
\]

**Theorem 7.6.** For every \( T > 0 \) and \( \kappa > 0 \),

\[
\lim_{\varepsilon \to 0} \sup_{\delta \geq A} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} |Z^\delta_t - \hat{Z}_t| \geq \kappa \right\} = 0.
\]

The proof will be given in Appendix C. Thus \( Z^{\delta,\varepsilon} \) converges to \( Z \) in probability in \( D_{\mathbb{R}}[0, \infty) \), and more precisely

\[
\left\langle M^{\delta,\varepsilon} \right\rangle_t \approx \frac{t}{T} - \frac{\varepsilon (\delta, \varepsilon) m t}{\varepsilon^2 T^2}.
\]

**Remark 7.7.** Let’s interpret (50). Let’s assume that the various constants (as introduced in the beginning of Subsection 7.1) are the same for the bid and ask queues. Then the correction to \( \left\langle M^{\delta,\varepsilon} \right\rangle \) is

\[
\frac{\varepsilon (\delta, \varepsilon) m}{\varepsilon^2 T^2} \leq \frac{\delta^2 T}{T^2} = \frac{\delta^2}{T^2} \left( \frac{\delta}{T} \right)^2 = \frac{q^2}{T^2}.
\]

This gives us (2).
Note that we have arranged things so that the HFT-er’s have a small effect compared to the dominant LT effects; as long as $\varsigma > 0$, the correction in (51) is small as $\varepsilon \searrow 0$. Roughly, the size and speed of changes in the midpoint price due to HFT-er’s are arranged to be negligible compared to the corresponding effects of LT-er’s. If $\varsigma$ of (24) is close to zero, then the correction of (51) will be larger; nevertheless, we are primarily interested in regimes where the HFT-er’s play a smaller role than the LT-er’s.

8. Conclusion

We have understood a simple and highly stylized model for the front of the order book, where the Level I orders are due to both low-frequency (liquidity) and high-frequency traders. We are interested in a regime where the liquidity traders are the primary ‘movers’ of the market and the high frequency traders play a secondary role. While there is an increasing interest in modelling the order book as a whole, we take as our starting point macroscopic diffusivity of the stock. The traders in our model are uninformed and non-strategic; the focus is the interaction between scales.

In the regime of interest, we find that the high-frequency traders can have a quantifiable effect on volatility. In an attempt to statistically model a range of high frequency traders, we allow the rates at which the high-frequency traders add and remove liquidity to fluctuate according to a CIR process. We further find that, in our model, there can be two regimes, depending on whether the dominant stochasticity in the order flow of the high-frequency traders is due to the Poisson stochasticity of the order flow or due to the CIR fluctuations of the order rate. Our analysis suggests that there should be something like an ‘HFT premium’ in volatility for stocks which are traded by HFT participants.

Appendix

A. Proofs for liquidity traders

We here want to prove the results of Section 3 concerning liquidity traders.

To begin, let’s show that the laws of the $\tau_{L,\delta}/\delta^2$’s are tight; this suggest that we have the correct scaling in (12).

**Lemma A.1.** For each $p \in (0, 1)$, we have that

$$\sup_{\delta \in (0, 1)} \mathbb{E} \left[ \frac{\tau_{L,\delta}^p}{\delta^2} \right] < \infty. \quad (52)$$

**Proof.** Define

$$Z_t \equiv \frac{Q_t^{\delta}}{\delta} - q \delta t + \delta^{\nu - \gamma_0} \ell t$$

$$= Q_t^{\delta} + \delta^{\nu - \gamma_0} \langle \ell - \hat{T} \rangle, \quad t \geq 0$$

Then $Z$ is a martingale with $Z_0 = 0$.

We now bound how unlikely it is that $\delta^{-2}\tau_{L,\delta}$ is much larger than $\hat{T}$. Fix $R > 0$. If $\delta^{-2}\tau_{L,\delta} > \hat{T}(1 + R)$, then $Q_{\delta^2\hat{T}(1 + R)}^{\delta} > \hat{T}(1 + R) > 0$. Thus

$$\mathbb{P}\left\{ \delta^{-2}\tau_{L,\delta} > \hat{T}(1 + R) \right\} \leq \mathbb{P}\left\{ Q_{\delta^2\hat{T}(1 + R)}^{\delta} > 0 \right\}$$

$$= \mathbb{P}\left\{ Z_{\delta^2\hat{T}(1 + R)} > \delta^2 \langle \ell - \hat{T} \rangle R \right\} \leq \frac{1}{\langle \ell^2 R^2 \rangle \delta^2 R^2}$$

Let’s now look at the evolution of $Z^2$. We have that

$$Z_t^2 = \int_{s=0}^{t} Y_s ds + M_t$$

where $M$ is a martingale with $M_0 = 0$ and where

$$Y_t = \delta^{\gamma_0} \left\{ (Z_t + \delta^{\nu_0})^2 - Z_t^2 \right\} \lambda_L^+$$

$$+ \delta^{\gamma_0} \left\{ (Z_t - \delta^{\nu_0})^2 - Z_t^2 \right\} \lambda_L^- - 2Z_t \delta^{\nu_0 - \gamma_0}$$

$$= \delta^{2\nu_0 - \gamma_0} \left\{ \lambda_L^+ \lambda_L^- \right\}. \quad \text{Thus}$$

$$\mathbb{E}[Z_t^2] = \left\{ \lambda_L^+ \lambda_L^- \right\} \delta^{2\nu_0 - \gamma_0} t$$

and hence (keeping (13)) in mind, we have that

$$\frac{\mathbb{E}\left[ Z_{\delta^2\hat{T}(1 + R)}^2 \right]}{\delta^{2\nu_0} R^2} \leq \left\{ \lambda_L^+ \lambda_L^- \right\} \hat{T} \frac{1 + R}{R^2} \delta^{2\nu_0 - \gamma_0 - 2}$$

$$= \left\{ \lambda_L^+ \lambda_L^- \right\} \hat{T} \frac{1 + R}{R^2} \delta^{\gamma_0 - 2}$$

$$\leq \left\{ \lambda_L^+ \lambda_L^- \right\} \hat{T} \frac{1 + R}{R^2}$$

for all $\delta \in (0, 1)$ and $R > 0$. \hfill \square
We can finally bound the expectation in (52). Fix \( p \in (0,1) \). We have that
\[
E \left[ \left| \frac{\tau_{\ell,\delta}^{(p)}}{\delta^2} \right|^p \right] = p \int_0^\infty s^{p-1} \mathbb{P} \left\{ \delta^{-2} \tau_{\ell,\delta} > s \right\} ds
\]
\[
\leq p \int_0^{2\hat{T}} s^{p-1} ds + p \int_{s=2T}^\infty s^{p-1} \mathbb{P}
\times \left\{ \delta^{-2} \tau_{\ell,\delta} > s \right\} ds
\leq (2\hat{T})^p + p\hat{T}^p \int_{R=1}^\infty (1+R)^{p-1} \mathbb{P}
\times \left\{ \delta^{-2} \tau_{\ell,\delta} > \hat{T}(1+R) \right\} dR
\leq (2\hat{T})^p + p(\lambda_L^+ + \lambda_L^-) \frac{\hat{T}^{p-1}}{\ell^2} \int_{R=1}^\infty \frac{(1+R)^p}{R^2} dR. \tag{53}
\] Since \( p < 1 \), the integral in this expression is finite, and the claim holds.

We now prove Lemma 3.1. To set the stage for the more detailed calculations associated with the HFT-er’s, our proof will not be the most direct. Namely, since \( \hat{T} \) is constant, we can prove (12) by showing weak convergence; we will use the Laplace transform to do this. Namely, it suffices to show that for every \( a \geq 0 \),
\[
\lim_{\delta \downarrow 0} \mathbb{E} \left[ \exp \left[ -a \frac{\tau_{\ell,\delta}^{(p)}}{\delta^2} \right] \right] = \exp \left[ -a\hat{T} \right]. \tag{54}
\]
To proceed, let’s define
\[
G(b) \overset{\text{def}}{=} (e^{-b} - 1) \lambda_L^+ + (e^b - 1) \lambda_L^-
\]
for all \( b \in \mathbb{R} \). Note that \( G(0) = 0 \) and that \( \dot{G}(0) = \ell > 0 \). By the implicit function theorem we can thus find a \( \eta > 0 \) such that \( G \) is invertible on \( I \overset{\text{def}}{=} G(-\eta, \eta) \). Defining \( G^{-1} \) as the inverse, we thus have that \( G^{-1}(0) = 0 \) and \( G^{-1} \) is strictly positive on \( G(0, \eta) \).

**Proof of Lemma 3.1.** The heart of the claim is (54). Lemma A.1 ensures that the laws of \( \delta^{-2} \tau_{\ell,\delta} \) are compact. Combined with (54) and the uniqueness of Laplace transforms, we as a result know that \( \delta^{-2} \tau_{\ell,\delta} \) converges weakly to \( \hat{T} \). Since weak convergence to a constant implies convergence in probability, the entire claim indeed follows from (54). Clearly (54) holds for \( a = 0 \), so let’s assume that \( a > 0 \). We now use (13) to see that there is a \( \delta \in (0,1) \) such that \( q\delta^{-2} \in I \) if \( \delta \in (0,\bar{\delta}] \). For \( \delta \in (0,\bar{\delta}] \), define
\[
b_{\delta}(a) \overset{\text{def}}{=} G^{-1}(a\delta^{-2}). \tag{55}
\]
We note that
\[
\lim_{\delta \downarrow 0} b_{\delta}(a) = \frac{a}{\ell} \tag{56}
\]
and that there is a constant \( K > 0 \) such that \( 0 < b_{\delta}(a) < K \) for \( \delta \in (0,\bar{\delta}] \).

For \( x \in \mathbb{R} \) and \( \delta \in (0,\bar{\delta}] \), define
\[
f_{\delta}(t, x) \overset{\text{def}}{=} \exp \left[ -b_{\delta}(a)\delta^{-2 \nu_0} x - \frac{a}{\delta^2} t \right]. \tag{57}
\]
We will prove (54) by using the stochastic process \( \{ f_{\delta}(t, Q_t^{L,\delta}); t \geq 0 \} \).

Since \( Q_t^{L,\delta} \) jumps in increments of \( \delta^{-\nu_0} \), we know that
\[
Q_t^{L,\delta} \geq -\delta^{-\nu_0}. \tag{58}
\]
if \( 0 \leq t \leq \tau_{\ell,\delta} \) and \( \delta \in (0,\bar{\delta}] \). Since \( Q_t^{L,\delta} \leq 0 \), we furthermore have that
\[
0 \geq Q_t^{L,\delta} \geq -\delta^{-\nu_0}. \tag{59}
\]
Combining assumption (13) and (58), we have that
\[
b_{\delta}(a)Q_t^{L,\delta} \delta^{-2 \nu_0} \geq -b_{\delta}(a)\delta^{-2} \geq -K
\]
if \( 0 \leq t \leq \tau_{\ell,\delta} \) and \( \delta \in (0,\bar{\delta}] \), so in fact
\[
0 \leq f_{\delta}(t, Q_t^{L,\delta}) \leq e^K \tag{60}
\]
if \( 0 \leq t \leq \tau_{\ell,\delta} \) and \( \delta \in (0,\bar{\delta}] \). We also have that
\[
\frac{\partial f_{\delta}}{\partial t}(t, Q_t^{L,\delta}) + (A_t^{L,\delta}) f_{\delta}(t, Q_t^{L,\delta})
\]
\[
= \left\{ \begin{array}{l}
\frac{a}{\delta^2} + \frac{1}{\delta^{\gamma_0}} \left( e^{-b_{\delta}(a)\delta^{-2}} - 1 \right) \lambda_L^+ \\
\frac{1}{\delta^{\gamma_0}} \left( e^{b_{\delta}(a)\delta^{-2}} - 1 \right) \lambda_L^- \end{array} \right\} f_{\delta}(t, Q_t^{L,\delta})
\]
\[
= \frac{1}{\delta^{\gamma_0}} \left( -a\delta^{-2} + G(b_{\delta}(a)\delta^{-2}) \right) \times f_{\delta}(t, Q_t^{L,\delta}) = 0
\]
for $0 \leq t \leq \tau^{L, \delta}$ and $\delta \in (0, \bar{\delta}]$. Thus \( \{f_\delta(\tau^{L, \delta} \wedge t, Q^{L, \delta}_{\tau^{L, \delta} \wedge t}) \; t \geq 0 \} \) is a martingale and hence

\[
\mathbb{E}\left[f_\delta(\tau^{L, \delta} \wedge t, Q^{L, \delta}_{\tau^{L, \delta} \wedge t})\right] = \exp\left[-b_\delta(a)q\delta^t + a\gamma \tau^{L, \delta}\right] = \exp\left[-b_\delta(a)q\right]
\]

for all $t \geq 0$.

Since \( \mathbb{P}\{\tau^{L, \delta} < \infty\} = 1 \), we \( \mathbb{P} \)-almost-surely have that

\[
\lim_{t \to \infty} \tau^{L, \delta} \wedge t = \tau^{L, \delta} \quad \text{and} \quad \lim_{t \to \infty} Q^{L, \delta}_{\tau^{L, \delta} \wedge t} = Q^{L, \delta}_{\tau^{L, \delta}}.
\]

Thus we can use dominated convergence (again recall (13)) on the first term of (61) and (56) on the second, and collecting things together we indeed get that

\[
\lim_{\delta \to 0} \mathbb{E} \left[ \exp \left[ -a\gamma \frac{\tau^{L, \delta}}{\delta^t} \right] \right] = \exp \left[ -a\gamma t \right].
\]

This gives us (54).

\[ \square \]

### B. Proofs for high frequency traders

We here want to prove the results of Section 4. We will use the calculations of Appendix A as guidance.

We first show tightness of the \( \xi_{\delta, \varepsilon} \)’s of (43); i.e., an analogue of Lemma A.1. Amongst other things, this will imply that we have the right scaling in (43).

**Lemma B.1.** For each $p > 0$ and $A \in \mathcal{D}(\nu, \gamma)$,

\[
\sup_{(\delta, \varepsilon) \in A'} \mathbb{E}[|\xi_{\delta, \varepsilon}|^p] < \infty
\]

for all $A'$ which are \( \mathcal{A} \)-good.

The proof will be somewhat similar to that of Lemma A.1. For future reference, define

\[
Z_{t}^{\delta, \varepsilon} \overset{\text{def}}{=} Q_{t}^{\delta, \varepsilon} - q\delta^t + \delta^{\nu_{\varepsilon} - \gamma} t t
\]

\[
= Q_{t}^{\delta, \varepsilon} + \delta^{\nu_{\varepsilon} - \gamma} t (t - \delta^2 T)
\]

Similarly to the proof of Lemma A.1, if \( |\xi_{\delta, \varepsilon}| \) is large, then \( Z \) should be large; we will precisely quantify this a bit later.

Let’s return to the Poisson characterizations of (5) and (20). To revisit our notation, \( N^{L, \pm} \) and \( N^{\pm} \) are all independent standard Poisson processes. We also have the two rate processes \( \lambda^{\pm, \varepsilon} \) as in (14). Define

\[
\hat{Q}_{t}^{\delta, \varepsilon} \overset{\text{def}}{=} \delta^{\nu_{\varepsilon}} N^{L, +} \left( \frac{1}{\delta \gamma} \lambda_{L}^{+} t \right) - \delta^{\nu_{\varepsilon}} N^{L, -} \left( \frac{1}{\delta \gamma} \lambda_{L}^{-} t \right) + \varepsilon^{\nu_{\varepsilon}} N^{+} \left( \frac{1}{\varepsilon \gamma} \int_{s=0}^{t} \lambda_{L}^{+} \gamma ds \right) - \varepsilon^{\nu_{\varepsilon}} N^{-} \left( \frac{1}{\varepsilon \gamma} \int_{s=0}^{t} \lambda_{L}^{-} \gamma ds \right).
\]

Then \( Q_{t}^{\delta, \varepsilon} \) and \( \hat{Q}_{t}^{\delta, \varepsilon} \) have the same laws; thus \( Z_{t}^{\delta, \varepsilon} \) has the same law as

\[
Z_{t}^{\delta, \varepsilon} \overset{\text{def}}{=} Q_{t}^{\delta, \varepsilon} - q\delta^t + \delta^{\nu_{\varepsilon} - \gamma} t t
\]

\[
= Q_{t}^{\delta, \varepsilon} + \delta^{\nu_{\varepsilon} - \gamma} t (t - \delta^2 T).
\]

\[ \square \]
Let’s now center things. Define

\[ \tilde{X}^{t,+}(t) \overset{\text{def}}{=} N^{t,+}(t) - t \]
and \[ \tilde{X}^{t,-}(t) \overset{\text{def}}{=} N^{t,-}(t) - t. \]

Then

\[
\tilde{Z}^{t,\varepsilon} = \delta_{t}^{\varepsilon} \tilde{X}^{t,+} \left( \frac{1}{\delta_{t}^{\varepsilon}} \lambda_{t}^{L} t \right) - \delta_{t}^{\varepsilon} \tilde{X}^{t,-} \left( \frac{1}{\delta_{t}^{\varepsilon}} \lambda_{t}^{L} t \right) + \varepsilon^\nu X^+ \left( \frac{1}{\varepsilon^\gamma} \int_{s=0}^{t} \lambda_{s}^{+} \varepsilon ds \right) - \varepsilon^\nu X^- \left( \frac{1}{\varepsilon^\gamma} \int_{s=0}^{t} \lambda_{s}^{-} \varepsilon ds \right) + \varepsilon^{\nu\gamma} \int_{s=0}^{t} \{ \lambda_{s}^+ - \lambda_{s}^- \} ds \]

\[
= \delta_{t}^{\varepsilon} \tilde{X}^{t,+} \left( \frac{1}{\delta_{t}^{\varepsilon}} \lambda_{t}^{L} t \right) - \delta_{t}^{\varepsilon} \tilde{X}^{t,-} \left( \frac{1}{\delta_{t}^{\varepsilon}} \lambda_{t}^{L} t \right) + \varepsilon^\nu X^+ \left( \frac{1}{\varepsilon^\gamma} \int_{s=0}^{t} \lambda_{s}^{+} \varepsilon ds \right) - \varepsilon^\nu X^- \left( \frac{1}{\varepsilon^\gamma} \int_{s=0}^{t} \lambda_{s}^{-} \varepsilon ds \right) + \varepsilon^{1/2+\nu\gamma} \frac{\sigma_{-}^{L}}{\alpha_{-}} \int_{s=0}^{t} \sqrt{\lambda_{s}^{+} \alpha_{-} dW_{s}^{-} - \varepsilon^{1/2+\nu\gamma}} \]

\[
\times \left\{ \lambda_{t}^{+} - \lambda_{0}^{+} \right\} + \varepsilon^{1/2+\nu\gamma} \frac{\sigma_{+}^{L}}{\alpha_{+}} \int_{s=0}^{t} \sqrt{\lambda_{s}^{-} \alpha_{+} dW_{s}^{+}} \left\{ \lambda_{t}^{+} - \lambda_{0}^{+} \right\} \]

(64)

where we have used (22). We will use this representation to bound the behavior of \( Z^{t,\varepsilon} \).

We need still another result. We will want to bound the behavior of centered Poisson processes.

**Lemma B.2.** Let \( N \) be a standard Poisson process. For each \( p \geq 1 \), there is a \( K_{p,B.2} > 0 \) such that

\[
\mathbb{E} \left[ \sup_{0 \leq s \leq t} |N(s) - s|^p \right] \leq K_{p} t^{p/2}
\]

for all \( t \geq 1 \).

**Proof.** Fix an integer \( \bar{p} \). Since the terms in the Taylor expansion of \( x \mapsto e^x \) are nonnegative if \( x \geq 0 \), we have that

\[
e^x \geq x^{2\bar{p}}/(2\bar{p})!
\]

for all \( x \geq 0 \). Symmetrizing the left-hand side, we get that

\[
x^{2\bar{p}} \leq (2\bar{p})! \{ e^x + e^{-x} \}
\]

for all \( x \in \mathbb{R} \). Thus

\[
\mathbb{E} \left[ \left| \frac{N(t) - t}{\sqrt{t}} \right|^{2\bar{p}} \right] \leq (2\bar{p})! \left\{ \mathbb{E} \left[ \exp \left( \frac{N(t) - t}{\sqrt{t}} \right) \right] \right\}^\bar{p} + \mathbb{E} \left[ \exp \left( - \frac{N(t) - t}{\sqrt{t}} \right) \right] \right\} \}
\]

We can explicitly calculate that

\[
\mathbb{E} \left[ \exp \left( \pm \frac{N(t) - t}{\sqrt{t}} \right) \right] \right\} = \mathbb{E} \left[ \exp \left( \pm \frac{1}{\sqrt{t}} \right) \right] \right\} \}
\]

\[
\times \exp \left[ \mp \frac{1}{\sqrt{t}} \right] \right\} \}
\]

\[
= \exp \left[ t \left\{ e^{\pm 1/\sqrt{t}} - 1 \right\} \exp \left[ \mp \frac{1}{\sqrt{t}} \right] \right\} \}
\]

\[
= \exp \left[ t \left\{ e^{\pm 1/\sqrt{t}} - 1 \mp 1/\sqrt{t} \right\} \right\} \}
\]

\[
= \exp \left[ G(\pm 1/\sqrt{t}) \right] \right\} \}
\]

where

\[
G(z) = \int_{s=0}^{t} (1 - s)e^{zs} ds
\]

for all \( z \in \mathbb{R} \). Note that \( |G(z)| \leq e/2 \) if \( |z| \leq 1 \). We thus get that if \( t \geq 1 \),

\[
\mathbb{E} \left[ \left| \frac{N(t) - t}{\sqrt{t}} \right|^{2\bar{p}} \right] \right\} \}
\]

where \( K \overset{\text{def}}{=} 2(2\bar{p})!e^{c/2} \). By Doob’s inequality, we can finish the proof; we have that
\[ \mathbb{E} \left[ \sup_{0 \leq t \leq T} |N(s) - s|^p \right] \leq \left( \frac{p}{p-1} \right)^p \mathbb{E} \left[ |N(t) - t|^{2p} \right] \]
\[
\leq \left( \frac{p}{p-1} \right)^p \left\{ \mathbb{E} \left[ |N(t) - t|^{2p} \right] \right\}^{\nu/(2p)}
\]
\[
= \left( \frac{p}{p-1} \right)^p \left\{ \mathbb{E} \left[ \left| N(t) - t \right|^{2p} \right] \right\}^{\nu/(2p)}
\]
\[
t^{p/2} \leq \left( \frac{p}{p-1} \right)^p \mathbb{E} \left[ |N(t) - t|^{2p} \right]^{p/(2p)}
\]

\[ \square \]

**Proof of Lemma B.1.** Fix \( \Xi > 0 \) and suppose that \( \delta \) and \( \varepsilon \) in \((0, 1)\) satisfy
\[
\varepsilon \leq (\gamma^A)^{1/2}/\delta \text{ or } \varepsilon \leq \delta^2 \Theta(\delta, \varepsilon, -1).
\]
(we will not explicitly bound \( \varepsilon^{(\gamma^A)^{1/2}}/\delta \) or \( \varepsilon^2/\delta^2\)). Fix now \( R > 0 \). We want to connect \( \{ |\xi_{\delta, \varepsilon}| \geq R \} \) to \( \delta^2 \Theta(\delta, \varepsilon, -1) \).

Assume that \( \xi_{\delta, \varepsilon} > -R \); i.e., \( \gamma_{\delta, \varepsilon} < -R \), and \( \delta^2 \Theta(\delta, \varepsilon, -1) \).

Thus there is a \( t^* \in (0, \delta^2 \Theta(\delta, \varepsilon, -1)) \) such that \( Q_{\delta, \varepsilon}^1 \leq 0 \). Consequently
\[
Z_{\delta, \varepsilon}^1 \leq \delta^2 \Theta(\delta, \varepsilon, -1) \leq -\delta^2 \ell R,
\]
Thus
\[
\sup_{0 \leq t \leq \delta^2 \Theta(\delta, \varepsilon, -1)} |Z_t| \geq \delta^2 \ell R.
\]

Alternately, assume that \( \xi_{\delta, \varepsilon} > R \), then \( Q_{\delta, \varepsilon}^1 \) has not touched 0 by time \( \delta^2 \Theta(\delta, \varepsilon, -1) \).

Thus \( Q_{\delta, \varepsilon}^2 > 0 \), so
\[
Z_{\delta^2 \Theta(\delta, \varepsilon, -1)} \leq \delta^2 \Theta(\delta, \varepsilon, -1) \leq \delta^2 \ell R;
\]
and hence
\[
|Z_{\delta^2 \Theta(\delta, \varepsilon, -1)}| \geq \delta^2 \ell R;
\]

Let’s collect things together. Under (65),
\[
\max \{ \delta^2 \Theta(\delta, \varepsilon, -1), \delta^2 \Theta(\delta, \varepsilon, R) \} \leq \delta^2 \Theta(1 + \Xi R/T),
\]

Thus
\[
\mathbb{P} \left\{ \{ \xi_{\delta, \varepsilon} \geq R \} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

We now use Markov’s inequality. Fix \( \bar{p} > \max \{1, 2p\} \).

Thus
\[
\mathbb{P} \left\{ \{ \xi_{\delta, \varepsilon} \geq R \} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

\[
\leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

where (using (64)),
\[
A_{1, \delta, \varepsilon}^i(R) \text{ def } \frac{1}{(\delta^2)^{p/R}} 
\]
\[
\times \left[ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} \left| \delta^2 \Theta(\delta, \varepsilon, -1) \right|^p \right]
\]
\[
\leq \mathbb{P} \left\{ \frac{1}{(\delta^2)^{p/R}} \right\}
\]

\[
\mathbb{P} \left\{ \{ \xi_{\delta, \varepsilon} \geq R \} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

\[
\leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

\[
\times \left[ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} \left| \delta^2 \Theta(\delta, \varepsilon, -1) \right|^p \right]
\]

\[
\leq \mathbb{P} \left\{ \frac{1}{(\delta^2)^{p/R}} \right\}
\]

\[
\mathbb{P} \left\{ \{ \xi_{\delta, \varepsilon} \geq R \} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

\[
\leq \mathbb{P} \left\{ \frac{1}{(\delta^2)^{p/R}} \right\}
\]

\[
\mathbb{P} \left\{ \{ \xi_{\delta, \varepsilon} \geq R \} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

\[
\leq \mathbb{P} \left\{ \frac{1}{(\delta^2)^{p/R}} \right\}
\]

\[
\mathbb{P} \left\{ \{ \xi_{\delta, \varepsilon} \geq R \} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq \delta^2 T(1 + \Xi R/T)} |Z_t^{\delta, \varepsilon}| \geq \delta^2 \ell R \right\}
\]

\[
\leq \mathbb{P} \left\{ \frac{1}{(\delta^2)^{p/R}} \right\}
\]
Let’s now use Lemma B.2. We have that
\[
A^\pm_{1,\delta,\epsilon}(R) \leq K_{\rho/2,B,2}^\pm \frac{\delta \nu_0}{(\delta \epsilon^2)^P R^p} \times \left( \frac{\lambda^{\pm}_P \delta^2 \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2} \\
= K_{\rho/2,B,2}^\pm \delta^{\rho(\nu - \gamma)/2} \quad \left( \frac{\lambda^{\pm}_P \delta \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2} \\
A^\pm_{2,\delta,\epsilon}(R) \leq K_{\rho/2,B,2}^\pm \frac{\delta \nu_0}{(\delta \epsilon^2)^P R^p} \mathbb{E} \times \left( \frac{1}{\epsilon^{\gamma}} \int_{s=0}^{T} \lambda^{\pm}_P \epsilon^{\rho/2} ds \right)^{\rho/2} \\
\leq K_{\rho/2,B,2}^\pm \delta^{\rho(\nu - \gamma)/2} \quad \left( \frac{\lambda^{\pm}_P \delta \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2} \\
\times \mathbb{E} \left[ \int_{s=0}^{T} \lambda^{\pm}_P \epsilon^{\rho/2} ds \right] \\
\leq K_{\rho/2,B,2}^\pm \delta^{\rho(\nu - \gamma)/2} \quad \left( \frac{\lambda^{\pm}_P \delta \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2} \\
\times \left( \frac{\lambda^{\pm}_P \delta \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2} \\
\leq K_{\rho/2,B,2}^\pm \delta^{\rho(\nu - \gamma)/2} \quad \left( \frac{\lambda^{\pm}_P \delta \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2} \\
\leq K_{\rho/2,B,2}^\pm \delta^{\rho(\nu - \gamma)/2} \quad \left( \frac{\lambda^{\pm}_P \delta \hat{T}(1 + \Xi R/T)}{\delta^{P+\rho}} \right)^{\rho/2}
\]

where \( K_{\rho/2,B}^\pm \) is an appropriate Burkholder constant.
Collecting things together, we see that there is a constant \( K \) such that
\[
\mathbb{P} \left\{ |\xi_{\delta,\epsilon}| \geq R \right\} \leq \frac{K}{R^{\rho/2}}
\]
for \( R \geq 1 \). Thus, as in (53),
\[
\mathbb{E} \left[ |\xi_{\delta,\epsilon}|^p \right] \\
= p \int_{s=0}^{T} s^{p-1} ds + p \int_{s=2T}^{\infty} s^{p-1} \mathbb{P} \left\{ |\xi_{\delta,\epsilon}| > s \right\} ds \\
\leq (2T)^p + p \int_{R=1}^{\infty} \hat{T}(1 + R)^{p-1} \mathbb{P} \left\{ |\xi_{\delta,\epsilon}| > \hat{T}(1 + R) \right\} dR \\
\leq (2T)^p + p K \hat{T}^p \int_{R=1}^{\infty} (1 + R)^{p-1} \mathbb{P} \left\{ |\xi_{\delta,\epsilon}| > \hat{T}(1 + R) \right\} dR.
\]
and since \( \hat{p}/2 > p \), the claim indeed follows. \( \square \)
We next want to prove Proposition 5.1. Equivalently, we want to show that
\[
\lim_{(\delta,\epsilon) \in A} \mathbb{E} \left[ \exp \left[ i \theta \xi_{\delta,\epsilon} \right] \right] = \exp \left[ -\frac{1}{2} \frac{\sigma^2}{\ell^2} \theta^2 \right] 
\]
for all \( \theta \in \mathbb{R} \). We face two major complications. Firstly, we here have two scales \( \delta, \epsilon \) instead of the single scale \( \delta \) of Lemma 3.1. Secondly, the stochasticity of the rates \( \lambda_{\delta,\epsilon} \) requires another averaging procedure.

Let’s first approximate the stochasticity of the order queue empties under the dynamics of (17). Define
\[
\hat{Q}_t^{\delta,\epsilon} \overset{\text{def}}{=} Q_t^{\delta,\epsilon} + \frac{\epsilon^{1+\nu-\gamma}}{\alpha_+} \left\{ \lambda_{t}^{\delta,\epsilon} - \lambda_{t}^{\bar{\delta},\epsilon} \right\} - \frac{\epsilon^{1+\nu-\gamma}}{\alpha_-} \left\{ \lambda_{t}^{-\delta,\epsilon} - \lambda_{t}^{-\bar{\delta},\epsilon} \right\}.
\]

**Remark B.3.** Let’s go back to the calculations of Remark 4.5. Instead of using the quadratic formula to solve the characteristic equation (31), we can use a recursive method. We can rewrite (31) as
\[
A = \frac{\theta}{b_0\sigma} + \frac{\sigma^2}{2b_0} A^2 = \frac{\theta}{b_0\sigma} \sqrt{b_0} + \frac{\sigma^2}{2b_0} A^2 \quad (66)
\]
(we have used here (34)). Let’s try to solve this recursively and to simultaneously identify the dominant parts. The approximation of (33) is that
\[
A \approx \frac{\theta \sqrt{b_0}}{x \sqrt{\sigma}} - \frac{1}{2x} \theta^2
\]
(where we have again used (34)). To capture this, let’s assume that
\[
A = \frac{\theta \sqrt{b_0}}{x \sqrt{\sigma}} + \frac{x^2}{2x} G = \frac{\theta \sqrt{b_0}}{x \sqrt{\sigma}} + \frac{x^2}{2x} b \quad (66)
\]
for some bounded \( G \). Inserting this in (66), we have that
\[
\frac{\theta \sqrt{b_0}}{x \sqrt{\sigma}} + \frac{x^2}{2x} G = \frac{\theta \sqrt{b_0}}{x \sqrt{\sigma}} + \frac{\sigma^2}{2b} \frac{x}{2x} b \left( \theta + \frac{\sigma}{\sqrt{b} \sigma} G \right)^2 
\]
which is equivalent to
\[
G = \frac{1}{2} \left( \theta + \frac{\sigma}{\sqrt{b} \sigma} G \right)^2 \quad (67)
\]
A refinement of these calculations will lay at the heart of our efforts to capture the dominant asymptotics of the statistics of the time at which the order queue empties under the dynamics of (17). Our arguments, however, are complicated by a number of scales. In fact, let’s complete our analysis of (67); it will provide a useful reference. Set \( G_0 = 0 \) and recursively define
\[
G_{n+1} = \frac{1}{2} \left( \theta + \frac{\sigma}{\sqrt{b} \sigma} G_n \right)^2.
\]
We claim that if \( \sigma/\sqrt{b} \) is sufficiently small, we get a contractive recursion. Fix \( K > \theta^2 \) and suppose that
\[
\frac{\sigma}{\sqrt{b} \sigma} \leq \min \left\{ 1, \sqrt{\frac{K - \theta^2}{K^2}}, \frac{1}{2(\theta + K)} \right\}.
\]
Suppose that \( |G_n| \leq K \) (which is indeed the case if \( n = 0 \)). Then
\[
|G_{n+1}| \leq \theta^2 + \frac{\sigma}{\sqrt{b} \sigma} |G_n|^2 \leq \theta^2 + (K - \theta^2) \leq K.
\]
Thus \( |G_n| \leq K \) for all \( n \in \mathbb{N} \). We then have that
\[
|G_{n+1} - G_n| \leq \frac{\sigma}{\sqrt{b} \sigma} |G_n - G_{n-1}| \leq \frac{2(\theta + \frac{\sigma}{\sqrt{b} \sigma} |G_n|)^2}{2} \leq \frac{\sigma}{\sqrt{b} \sigma} |G_n - G_{n-1}| \leq \frac{1}{2} |G_n - G_{n-1}|
\]
Thus the \( G_n \)’s must converge to a limit.

Recalling our calculations of Appendix A, let’s construct the analogue of (57). In place of the function \( b_0 \) of (55), we need to capture both the fluid limits of the liquidity traders and the heavy traffic approximation of the high-frequency trader. We begin by something like an implicit function theorem which focuses on these dominant effects. Define
\[
\mathcal{E}_2(z) \overset{\text{def}}{=} e^z - \{1 + z\} \quad \text{and} \quad \mathcal{E}_3(z) \overset{\text{def}}{=} e^z - \left\{1 + z + \frac{1}{2} z^2\right\}
\]
for all \( z \in \mathbb{C} \). Then define
\[
\tilde{E}_{\delta, \varepsilon}(z) = \frac{1}{\delta^{\gamma_0}} \left\{ E_2(\delta^{2\nu_0} z) \lambda_L^+ + E_2(-\delta^{2\nu_0} z) \lambda_L^- \right\} \\
+ \frac{1}{\varepsilon} \left\{ E_3(\varepsilon^2 z) + E_3(-\varepsilon^2 z) \right\} \lambda
\]
for all \( \delta \) and \( \varepsilon \) in \((0, 1)\) and all \( z \in \mathbb{R} \).

**Lemma B.4.** There is a \( K_{B,4} > 0 \) such that for \( \delta \) and \( \varepsilon \) in \((0, 1)\)
\[
|\tilde{E}_{\delta, \varepsilon}(z)| \leq K_{B,4} \left\{ \delta^{2\nu_0 - \gamma_0} + \varepsilon^{2\nu_0 - \gamma_0} \right\} |z|^2 \\
\times \exp \left\{ |\delta^{\nu_0} + \varepsilon^{\nu_0}| |\Re(z)| \right\}
\]
\[
|\tilde{E}_{\delta, \varepsilon}(z)| \leq K_{B,4} \left\{ \delta^{2\nu_0 - \gamma_0} + \varepsilon^{2\nu_0 - \gamma_0} \right\} |z| \left( |z_1| + |z_2| \right) \\
\times \exp \left\{ |\delta^{\nu_0} + \varepsilon^{\nu_0}| |\Re(z_1)| \right\}
\]
\[
|E_3(z)| \leq K_{B,4} |z|^2 \exp \left\{ |\Re(z)| \right\}
\]
for all \( z \in \mathbb{C} \) (where, as usual, \( z = \Re(z) + i\Im(z) \) for all \( z \in \mathbb{C} \)) and for \( z_1 \) and \( z_2 \) in \( \mathbb{C} \).

The first bound on \( \tilde{E}_{\delta, \varepsilon} \) will be used to show an existence result (Proposition B.5); the second will be used to show the asymptotics of Proposition 5.1 (the proof of which is at the end of this section).

**Proof of Lemma B.4.** Note that
\[
|e^z| \leq e^{\Re(z)} \quad \text{and} \quad 1 \leq e^{\Re(z)} \tag{68}
\]
for all \( z \in \mathbb{C} \).

To warm up, let’s prove the stated bound on \( E_3 \). Fix \( z \in \mathbb{C} \) and define
\[
f(s) \overset{\text{def}}{=} e^{sz} - \left\{ 1 + sz + \frac{1}{2} (sz)^2 \right\}.
\]
Then \( f(0) = \tilde{f}(0) = 0 \), and
\[
E_3(z) = f(1) = \int_{s=0}^{1} (1-s) \tilde{f}(s) ds
\]
\[
= z^2 \int_{s=0}^{1} (1-s) \left\{ e^{sz} - 1 \right\} ds.
\]
The claimed bound on \( E_3 \) follows (use (68)).

We now turn to bounding \( E_{\delta, \varepsilon} \). Fix \( z \in \mathbb{C} \).

\[
2f_2(s) \overset{\text{def}}{=} e^{sz} - 1 - sz
\]
and
\[
f_3(s) \overset{\text{def}}{=} e^{sz} + e^{-sz} - 2 - (sz)^2
\]
for all \( s \in [0, 1] \). First note that \( f_2(0) = \tilde{f}_2(0) = 0 \) so
\[
E_2(z) = f_2(1) = \int_{s=0}^{1} (1-s) f_2^{(2)}(s) ds
\]
\[
= z^2 \int_{s=0}^{1} (1-s) e^{sz} ds.
\]
Thus
\[
|E_2(z)| \leq |z|^2 \int_{s=0}^{1} (1-s) e^{sz} ds \leq \frac{|z|^2}{2} e^{\Re(z)}
\]
for all \( z \in \mathbb{C} \). Next note that \( f_3^{(i)}(0) = 0 \) for \( i \in \{0, 1, 2, 3\} \). Using the Taylor expansions of degree 2 and 3, we see that
\[
E_3(z) + E_3(-z)
\]
\[
= f_3(1) = \int_{s=0}^{1} (1-s) f_3^{(2)}(s) ds
\]
\[
= z^2 \int_{s=0}^{1} (1-s) \left( e^{sz} + e^{-sz} - 2 \right) ds
\]
\[
E_3(z) + E_3(-z)
\]
\[
= f_3(1) = \frac{1}{3!} \int_{s=0}^{1} (1-s)^3 f_3^{(4)}(s) ds
\]
\[
= \frac{z^4}{6} \int_{s=0}^{1} (1-s)^3 \left( e^{sz} + e^{-sz} \right) ds
\]
for all \( z \in \mathbb{C} \). Consequently
\[
|E_3(z) + E_3(-z)|
\]
\[
\leq |z|^2 \int_{s=0}^{1} (1-s) \left\{ |e^{sz}| + |e^{-sz}| + 2 \right\} ds \leq 2|z|^2 e^{\Re(z)}
\]
\[
|E_3(z) + E_3(-z)|
\]
\[
\leq \frac{|z|^4}{6} \int_{s=0}^{1} (1-s)^3 \left\{ |e^{sz}| + |e^{-sz}| \right\} ds \leq \frac{|z|^4}{12} e^{\Re(z)}
\]
for all $z \in \mathbb{C}$. Collecting things together and using the first bound of (68), we get that

$$|\mathcal{E}_{\delta,x}(z)| \leq \frac{\lambda_1^x + \lambda_2^x}{2} \delta^{2\nu_x - \gamma_x} |z|^2 \exp[\delta^{\nu_x} |\mathcal{R}(z)|] + 2 |\lambda| e^{2\nu_x - \gamma}|z|^2 \exp[\delta^{\nu_x} |\mathcal{R}(z)|]$$

for all $z \in \mathbb{C}$. The bounds on the size of $\mathcal{E}_{\delta,x}$ follow.

To see the claimed continuity, fix $z_1$ and $z_2$ in $\mathbb{C}$ Define

$$f_2(s) \overset{\text{def}}{=} \exp[z_1 + s(z_2 - z_1)] - 1 - (z_1 + s(z_2 - z_2))$$

$$f_3(s) \overset{\text{def}}{=} \exp[z_1 + s(z_2 - z_1)] + \exp[-z_1 - s(z_2 - z_1)] - 2 - (z_1 + s(z_2 - z_2))^2$$

Then

$$\mathcal{E}(z_2) - \mathcal{E}(z_1) = f_2(1) - f_2(0)$$

$$= (z_2 - z_1) \int_{s=0}^{1} \left\{ \exp[z_1 + s(z_2 - z_1)] - 1 \right\} \, ds$$

$$= (z_2 - z_1) \int_{s=0}^{1} \left\{ z_1 + s(z_2 - z_1) \right\} \int_{r=0}^{1} \exp[r(z_1 + s(z_2 - z_1))] \, dr \, ds$$

$$= (z_2 - z_1) \int_{s=0}^{1} \left\{ (1 - s)z_1 + z_2s \right\} \int_{r=0}^{1} \exp[r(1 - s)z_1 + rsz_2] \, dr \, ds$$

and

$$\{\mathcal{E}_3(z_2) + \mathcal{E}_3(-z_2)\} - \{\mathcal{E}_3(z_1) + \mathcal{E}_3(z_1)\}$$

$$= f_3(1) - f_3(0) = (z_2 - z_1) \int_{s=0}^{1} \{\exp[z_1 + s(z_2 - z_1)] - \exp[-z_1 - s(z_2 - z_1)] - 2(z_1 + s(z_2 - z_1))\} \, ds$$

$$= (z_2 - z_1) \int_{s=0}^{1} (z_1 + s(z_2 - z_1)) \int_{r=0}^{1} \exp[r(z_1 + s(z_2 - z_1))] \, dr \, ds$$

Thus

$$|\mathcal{E}_3(z_2) - \mathcal{E}_3(z_1)| \leq |z_2 - z_1| \left\{ |z_1| + |z_2| \right\} \left\{ e^{\alpha z_1 + |\mathcal{R}(z)|} + e^{\alpha z_2 + |\mathcal{R}(z)|} \right\}$$

$$\leq 2|z_2 - z_1| \left\{ |z_1| + |z_2| \right\} \left\{ e^{\alpha z_1 + |\mathcal{R}(z)|} + e^{\alpha z_2 + |\mathcal{R}(z)|} \right\}$$

(we have used Jensen’s inequality to see that $e^{\alpha x + (1 - \alpha)y} \leq \alpha e^x + (1 - \alpha)e^y$ for all $x$ and $y$ in $\mathbb{R}$ and $\alpha \in [0, 1]$). Combine things together to get the claimed continuity on $\mathcal{E}_{\delta,x}$; we here use both inequalities in (68). □

Let’s now develop an analogue of (66). First of all let’s note that

$$\mathcal{E}^\nu = \frac{e^{\nu - \gamma/2}}{\delta ^{1 - \nu_x + \gamma_x/2}} \mathcal{E}^\gamma/\delta \leq \left( \frac{e^{\nu - \gamma/2}}{\delta ^{1 - \nu_x + \gamma_x/2}} \right) \left( \frac{e^{\gamma/2 \wedge 1}}{\delta} \right).$$

and also that since $\zeta \leq \nu - \gamma/2$ and $\zeta \leq \frac{1}{2} + \nu - \gamma$,

$$\frac{e^{\nu - \gamma/2}}{\zeta} \leq 1 \quad \text{and} \quad \frac{e^{1/2 + \nu - \gamma}}{\zeta} \leq 1$$

for $\zeta \in (0, 1)$.

Let’s next fix $\theta \in \mathbb{R}$ and set

$$m_1 \overset{\text{def}}{=} \frac{|\theta|}{\ell} \quad \text{and} \quad m_2 \overset{\text{def}}{=} \frac{\sigma^2 \chi^2}{2 \ell}.$$

(where $\chi$ was defined in (26)).

**Proposition B.5.** Define

$$I \overset{\text{def}}{=} m_1^2 \left( m_2 + \frac{K_{B,A}}{\ell} \right) + 1.$$  

(71)
Then there is a $\delta_0 > 0$ such that if $\delta$ and $\varepsilon$ in $(0, 1)$ are such that
\[
\delta \leq \delta_0, \quad \frac{\varepsilon^{\gamma/2\Lambda_1}}{\delta} \leq \delta_0, \quad \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}/2} \geq \frac{1}{\delta_0},
\]
and \[
\frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} \leq \delta_0,
\]
then there is a $G_{\delta,\varepsilon} \in \mathbb{C}$ with $|G_{\delta,\varepsilon}| \leq F$ such that
\[
A_{\delta,\varepsilon} \overset{\text{def}}{=} \frac{i\theta}{\ell \delta^\varepsilon} + \frac{G_{\delta,\varepsilon}}{\delta^\varepsilon} = \frac{1}{\delta^\varepsilon} \left( \frac{i\theta}{\ell} + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} G_{\delta,\varepsilon} \right)\]
satisfies
\[
A_{\delta,\varepsilon} = \frac{i\theta}{\ell \delta^\varepsilon} + \frac{\varepsilon^\lambda}{\delta^{\nu_0-\gamma_0}} \left( \frac{\sigma^2 \chi^2 \Lambda_1}{2\ell} \right) A_{\delta,\varepsilon}^2 + \frac{1}{\delta^\varepsilon} \tilde{\varepsilon}_{\delta,\varepsilon} (A_{\delta,\varepsilon}).
\]
(73)

Compare (72) to (40).

Proof of Proposition B.5. As much as possible, we would like to follow the arguments of Remark B.3 with scales as in (36). Inserting (73) into (74), we see that $G_{\delta,\varepsilon}$ should satisfy
\[
\frac{i\theta}{\ell \delta^\varepsilon} + \frac{1}{\delta^\varepsilon} G_{\delta,\varepsilon} = \frac{i\theta}{\ell \delta^\varepsilon} + \frac{\varepsilon^\lambda}{\delta^{\nu_0-\gamma_0}} \left( \frac{\sigma^2 \chi^2 \Lambda_1}{2\ell} \right)^2
\]
\[
+ \frac{\varepsilon^\lambda}{\delta^{\nu_0-\gamma_0}} \left( \frac{\sigma^2 \chi^2 \Lambda_1}{2\ell} \right)^2 A_{\delta,\varepsilon} + \frac{1}{\delta^\varepsilon} \tilde{\varepsilon}_{\delta,\varepsilon} (A_{\delta,\varepsilon})
\]
or equivalently
\[
G_{\delta,\varepsilon} = \left( \frac{\sigma^2 \chi^2 \Lambda_1}{2\ell} \right) \left( \frac{i\theta}{\ell} + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} G_{\delta,\varepsilon} \right)^2
\]
\[
+ \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} \left( \frac{i\theta}{\ell} + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} G_{\delta,\varepsilon} \right)
\]
We want to solve this iteratively. Set $G_{\delta,\varepsilon}^0 = 0$ and recursively define
\[
G_{\delta,\varepsilon}^{n+1} = \left( \frac{\sigma^2 \chi^2 \Lambda_1}{2\ell} \right) \left( \frac{i\theta}{\ell} + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} G_{\delta,\varepsilon}^n \right)^2
\]
\[
+ \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} \left( \frac{i\theta}{\ell} + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} G_{\delta,\varepsilon}^n \right)
\]
Let’s first show that $|G_{\delta,\varepsilon}^n| \leq F$ for all $n \in \mathbb{N}$. Suppose that $\delta_0 > 0$ is such that
\[
|m_1 + \delta_0 F|^2
\]
\[
\times \left\{ m_2 + \frac{K_B \delta_0 (\delta_0^2 + 1)}{\ell} \exp \left[ \frac{\left( \delta_0^2 + \delta_0^\gamma - 2 \right)}{\delta_0^\varepsilon} F \right] \right\} \leq F;
\]
this is possible by our choice (71) of $F$. Suppose also that
\[
\delta \leq \delta_0, \quad \frac{\varepsilon^{\gamma/2\Lambda_1}}{\delta} \leq \delta_0, \quad \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}/2} \geq \frac{1}{\delta_0},
\]
and \[
\frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} \leq \delta_0.
\]
Suppose that $|G_{\delta,\varepsilon}^n| \leq F$ (which is true for $n = 0$). Using Lemma B.4,
\[
|G_{\delta,\varepsilon}^{n+1}|
\]
\[
\leq m_2 \left| m_1 + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} F \right|^2
\]
\[
+ K_B \frac{\delta_0^2 (\delta_0^2 + \varepsilon^\lambda - \varepsilon^\gamma)}{\ell (\delta_0^\varepsilon)^2}
\]
\[
\times \left\{ m_2 + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} F \right\} \left( \frac{\delta_0^\gamma + \delta_0^\varepsilon}{\delta_0^\varepsilon} \right)
\]
\[
\leq \left\{ m_2 + \frac{K_B \delta_0 (\delta_0^2 + 1)}{\ell} \left( \frac{\delta_0^\gamma + \delta_0^\varepsilon}{\delta_0^\varepsilon} \right)^2 \right\}
\]
\[
\times \left\{ m_2 + \frac{K_B \delta_0 (\delta_0^2 + 1)}{\ell} \left( \frac{\delta_0^\gamma + \delta_0^\varepsilon}{\delta_0^\varepsilon} \right)^2 \right\}
\]
\[
\times \left\{ m_2 + \frac{K_B \delta_0 (\delta_0^2 + 1)}{\ell} \left( \frac{\delta_0^\gamma + \delta_0^\varepsilon}{\delta_0^\varepsilon} \right)^2 \right\}
\]
\[
\leq F.
\]
Several auxiliary calculations have been used. Thanks to (13), $\delta \mapsto \delta^{\gamma-2}$ is nondecreasing on $(0, \infty)$. We have also used (69) and (70). Thus $|G_{\delta,\varepsilon}^n| \leq F$ for all $n \in \mathbb{N}$.

We next want to show that the $G_{\delta,\varepsilon}$’s converge. For $n \in \{2, 3, \ldots \}$, we have that
\[
G_{\delta,\varepsilon}^{n+1} - G_{\delta,\varepsilon}^n
\]
\[
= \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} \left( \frac{\sigma^2 \chi^2 \Lambda_1}{2\ell} \right) \left( \frac{i\theta}{\ell} + \frac{\varepsilon^\lambda}{\delta^{1+\nu_0-\gamma_0}} G_{\delta,\varepsilon}^n \right).
\]
Also, suppose that

\[
\|\mathcal{F} \leq +1
\]

and (72) holds, then

\[
A_{\delta,\varepsilon} = \delta \{ \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell} \right) + \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell^3} \right) \theta^2 \}
\]

Note that

\[
\left| \frac{1}{\delta^2} \left( \frac{\varepsilon^2}{\delta^1 + \nu_0 - \gamma_0} G_{\delta,\varepsilon}^n \right) \right| \leq \frac{1}{\delta^1} \left( G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right).
\]

Suppose that \( G_{\delta,\varepsilon}^n \leq F \) and \( G_{\delta,\varepsilon}^{n-1} \leq F \).

Let \( \Xi \in (0,\Xi_c) \) such that

\[
\Xi \left| m_1 + \Xi_c F \right| \left\{ 2m_2 + K_{B,4} (\Xi^2 + 1) \right\} \leq \frac{1}{2}.
\]

Also, suppose that \( \delta \) and \( \varepsilon \) in \((0,1)\) satisfy (72). We then have that

\[
|G_{\delta,\varepsilon}^{n+1} - G_{\delta,\varepsilon}^n| \leq \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} \left| m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right| \left( G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right)
\]

\[
+ K_{B,4} \left( \frac{2 \varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right) m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F
\]

\[
\times \left| G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right| \exp \left[ \frac{\varepsilon^5 + \varepsilon^\nu}{\delta^1 + \nu_0 - \gamma_0} F \right]
\]

\[
\leq \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} \left| m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right|
\]

\[
\times \left( 2m_2 + K_{B,4} \left( \frac{\varepsilon^5 - \gamma_0}{\delta^1 + \nu_0 - \gamma_0} + \frac{\varepsilon^5 - \gamma_0}{\delta^5} \right)^2 \right) \exp \left[ \frac{\varepsilon^5 - \gamma_0}{\delta^1 + \nu_0 - \gamma_0} F \right]
\]

\[
\times \left( G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right)
\]

\[
\leq \frac{1}{2} \left| G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right|
\]

We have again used (69) and (70). The existence of the claimed \( G_{\delta,\varepsilon} \) follows. 

Let’s rewrite the expansion of \( A_{\delta,\varepsilon} \) is an alternate way.

Define

\[
\tilde{G}_{\delta,\varepsilon} \equiv \delta \left\{ \frac{\varepsilon^2 \chi^2 \lambda}{\delta^1 + \nu_0 - \gamma_0} \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell} \right) + \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell^3} \right) \theta^2 \right\}
\]

\[
\quad + \frac{1}{\delta^1 + \nu_0 - \gamma_0} \tilde{G}_{\delta,\varepsilon}(\delta,\varepsilon) + \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell^3} \right) \theta^2
\]

\[
= \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell} \right) \left( \delta \tilde{G}_{\delta,\varepsilon}(\delta,\varepsilon) \right)^2 - \left( \frac{\varepsilon^2 \chi^2 \lambda}{2\ell} \right) \theta^2
\]

\[
+ \delta^2 \left( \frac{\varepsilon^2 \chi^2 \lambda}{2\ell^3} \right) \theta^2
\]

so that

\[
A_{\delta,\varepsilon} = \delta \left\{ \left( \frac{\sigma^2 \chi^2 \lambda}{2\ell^3} \right) \theta^2 + \tilde{G}_{\delta,\varepsilon} \right\} \quad (75)
\]

Note that if \( |G_{\delta,\varepsilon}| \leq F \) and (72) holds, then

\[
\left| \tilde{G}_{\delta,\varepsilon} \right| \leq \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} \left( \frac{\varepsilon^5 - \gamma_0}{\delta^1 + \nu_0 - \gamma_0} \right)^2 \exp \left[ \frac{\varepsilon^5 - \gamma_0}{\delta^1 + \nu_0 - \gamma_0} F \right]
\]

\[
\times \left( 2m_2 + K_{B,4} \left( \frac{\varepsilon^5 - \gamma_0}{\delta^1 + \nu_0 - \gamma_0} \right) m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right) \left| G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right|
\]

\[
\leq \left( \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} \right)^2 \left| m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right| \left( G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right)
\]

\[
+ \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} \left( m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right)
\]

\[
\times \left( 2m_2 + K_{B,4} \left( \frac{\varepsilon^5 - \gamma_0}{\delta^1 + \nu_0 - \gamma_0} \right) m_1 + \frac{\varepsilon^5}{\delta^1 + \nu_0 - \gamma_0} F \right) \left| G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right|
\]

\[
\leq \frac{1}{2} \left| G_{\delta,\varepsilon}^n - G_{\delta,\varepsilon}^{n-1} \right|
\]
Lemma B.4. We have also used (69) and (70). Thus

\[
\lim_{(\delta,\varepsilon) \to (0,0)} E_{\delta,\varepsilon} = 0. \tag{76}
\]

Define now

\[
X_t \overset{\text{df}}{=} \exp \left[ A_{\delta,\varepsilon}^t Q^{\delta,\varepsilon} + i \theta \left( \frac{t - T^2}{\delta^{1+\nu} + \gamma \varepsilon} \right) \right].
\]

We carefully write down the evolution of \(X\). We have that

\[
X_t = X_0 + \int_{s=0}^t Y_s X_s ds + M_t
\]

where \(M\) is a martingale and

\[
Y_t = \frac{1}{\delta^{\nu_0}} \left\{ \exp \left[ \nu_0 A_{\delta,\varepsilon}^t \right] - 1 \right\} \lambda_L^+ + \frac{1}{\delta^{\nu_0}} \left\{ \exp \left[ -\nu_0 A_{\delta,\varepsilon}^t \right] - 1 \right\} \lambda_L^- + \frac{1}{\varepsilon} \left\{ \exp \left[ e^{\nu_0^2} A_{\delta,\varepsilon}^t \right] - 1 \right\} \lambda_t^{+\varepsilon} + \frac{1}{\varepsilon} \left\{ \exp \left[ -e^{\nu_0^2} A_{\delta,\varepsilon}^t \right] - 1 \right\} \lambda_t^{-\varepsilon} - e^{\nu_0^2 - \gamma} A_{\delta,\varepsilon}^t (\lambda_L^{+\varepsilon} - \lambda_L^-) + e^{\nu_0 - \gamma} A_{\delta,\varepsilon}^t (\lambda_t^{-\varepsilon} - \lambda_t^+)/\lambda_L^+ + \frac{1}{\varepsilon} \left[ 2^{1+2(\nu_0 - \gamma)} A_{\delta,\varepsilon}^t \frac{\sigma^2}{\alpha_+^{+\varepsilon}} \lambda_t^{+\varepsilon} + \frac{e^{2^0 - \gamma} A_{\delta,\varepsilon}^t \frac{\sigma^2}{\alpha_-^{-\varepsilon}} \lambda_t^{-\varepsilon} + \frac{i \theta}{\delta^{1+\nu_0} + \gamma \varepsilon} \right. \]

At this point we see why \(A_{\delta,\varepsilon}\) was chosen to solve (74). We get that in fact

\[
Y_t = a_{\delta,\varepsilon,+} \left\{ \lambda_L^{+\varepsilon} - \lambda \right\} + a_{\delta,\varepsilon,-} \left\{ \lambda_t^{-\varepsilon} - \lambda \right\}
\]

where

\[
a_{\delta,\varepsilon,+} = \frac{1}{2} \frac{e^{2^0 - \gamma} A_{\delta,\varepsilon}^t \frac{\sigma^2}{\alpha_-^{-\varepsilon}} \lambda_t^{-\varepsilon} + \frac{e^{2^0 - \gamma} A_{\delta,\varepsilon}^t \frac{\sigma^2}{\alpha_-^{-\varepsilon}} \lambda_t^{-\varepsilon} + \frac{i \theta}{\delta^{1+\nu_0} + \gamma \varepsilon} }{\lambda_L^+} \]

In other words,

\[
X_t = X_0 + a_{\delta,\varepsilon,+} \int_{s=0}^t \left\{ \lambda_L^{+\varepsilon} - \lambda \right\} X_s ds + a_{\delta,\varepsilon,-} \int_{s=0}^t \left\{ \lambda_t^{-\varepsilon} - \lambda \right\} X_s ds + M_t
\]

where \(M\) is a martingale. To bound the error between \(\lambda^{\pm\varepsilon}\) and \(\lambda\), define

\[
\tilde{X}_t = X_t + \frac{\varepsilon}{\alpha_+} a_{\delta,\varepsilon,+} X_t \left\{ \lambda_L^{+\varepsilon} - \lambda_L^- \right\} + \frac{e^{2^0 - \gamma} A_{\delta,\varepsilon}^t \frac{\sigma^2}{\alpha_-^{-\varepsilon}} \lambda_t^{-\varepsilon} + \frac{i \theta}{\delta^{1+\nu_0} + \gamma \varepsilon} \}
\]

Then

\[
\tilde{X}_t = X_0 + \int_{s=0}^t \tilde{Y}_s X_s ds + M_t
\]
where $M$ is a martingale and
\[
\begin{align*}
\tilde{Y}_t &= \left\{ \frac{\varepsilon}{\alpha_+} a_{\delta,\varepsilon, t} \left( \lambda_t^+ - \lambda_t^- \right) - \frac{\varepsilon}{\alpha_-} a_{\delta,\varepsilon, -} \left( \lambda_t^- - \lambda_t^- \right) \right\} \\
&\quad \times \left\{ a_{\delta,\varepsilon, +} (\lambda_t^+ - \lambda_t^+) + a_{\delta,\varepsilon, -} (\lambda_t^- - \lambda_t^-) \right\} \\
&\quad + \frac{\sigma^2}{\alpha_+^2} \varepsilon^{1+\nu-\gamma} a_{\delta,\varepsilon, +} \lambda_t^{+\nu} \\
&\quad + \frac{\sigma^2}{\alpha_-^2} \varepsilon^{1+\nu-\gamma} a_{\delta,\varepsilon, -} \lambda_t^{-\nu}.
\end{align*}
\]

We can now prove the claimed asymptotics of $\tau^{\delta,\varepsilon}$. 

**Proof of Proposition 5.1.** First of all, the result is obvious if $\theta = 0$; we thus assume that $|\theta| > 0$. Let's also assume that $\delta$ and $\varepsilon$ are as in (72), so that $G_{\delta,\varepsilon}$ of Proposition (72) is well-defined and $|G_{\delta,\varepsilon}| \leq T$ where $T$ is as in (71).

Let's start by replacing $\tau^{\delta,\varepsilon}$ in the definition (43) of $\xi_{\delta,\varepsilon}$ by a ‘nicer’ stopping time. For $\delta$ and $\varepsilon$ in $(0, 1)$ and $L > 0$, define
\[
\begin{align*}
\theta_{\delta,\varepsilon,L} &= \inf \left\{ t \geq 0 : |\lambda_t^{+\varepsilon}| \geq \left( \frac{L T^{\delta/\varepsilon}}{\varepsilon} \right)^{1/4} \right\} \\
\text{or} \\
|\lambda_t^{-\varepsilon}| &\geq \left( \frac{L T^{\delta/\varepsilon}}{\varepsilon} \right)^{1/4} 
\end{align*}
\]
and $\tilde{\tau}^{\delta,\varepsilon}_L \overset{def}{=} \tau^{\delta,\varepsilon} \wedge \theta_{\delta,\varepsilon,L} \wedge (L T^{\delta/\varepsilon})$. We then have that
\[
\mathbb{E} \left[ \exp \left[ i\theta \xi_{\delta,\varepsilon} \right] \right] = \exp \left[ -\frac{\sigma^2 \lambda_t q}{2 \varepsilon^3} \right] = \sum_{i=1}^{3} \mathcal{E}^{(i)}(\delta, \varepsilon, L)
\]
where
\[
\begin{align*}
\mathcal{E}^{(1)}(\delta, \varepsilon, L) &= \mathbb{E} \left[ \exp \left[ i\theta \xi_{\delta,\varepsilon} \right] \right] - \mathbb{E} \left[ \tilde{X}_{\delta,\varepsilon} \right] \\
\mathcal{E}^{(2)}(\delta, \varepsilon, L) &= \mathbb{E} \left[ \tilde{X}_{\delta,\varepsilon} \right] - \tilde{X}_0 \\
\mathcal{E}^{(3)}(\delta, \varepsilon, L) &= \tilde{X}_0 - \exp \left[ -\sigma^2 \lambda_t q / 2 \varepsilon^3 \theta^2 \right]
\end{align*}
\]
Let's start at the end. We calculate that
\[
\mathcal{E}^{(3)}(\delta, \varepsilon, L) = \exp \left[ A_{\delta,\varepsilon} \theta^2 - \frac{i\theta T \delta^2}{\delta^{1-\nu} + \gamma / \varepsilon} \right]
\]
We have from (75) that
\[
A_{\delta,\varepsilon} \theta^2 - \frac{i\theta T \delta^2}{\delta^{1-\nu} + \gamma / \varepsilon} = -q \left( \frac{\sigma^2 \lambda_t}{2 \varepsilon^3} \right) \theta^2 + \tilde{G}_{\delta,\varepsilon} q
\]
and from (76) that
\[
\lim_{A \in D(\nu, \gamma)} A_{\delta,\varepsilon} \theta^2 - \frac{i\theta T \delta^2}{\delta^{1-\nu} + \gamma / \varepsilon} = -q \left( \frac{\sigma^2 \lambda_t}{2 \varepsilon^3} \right) \theta^2
\]
(recall that $\lim_{\varepsilon \downarrow 0} \chi_{\varepsilon} = 1$). Thus we indeed have that
\[
\lim_{A \in D(\nu, \gamma)} \mathcal{E}^{(3)}(\delta, \varepsilon, L) = 0
\]
for all $L > 0$.

Fix next $t \in [0, \tilde{\tau}^{\delta,\varepsilon}]$. Since the jumps of $Q$ are either $\delta^\nu$ or $\varepsilon^\nu$, we have that
\[
\tilde{Q}_{\delta,\varepsilon}^{\delta,\varepsilon} \geq -\delta^\nu - \varepsilon^\nu - 2^{1+\nu-\gamma} (L T^{\delta/\varepsilon})^{1/4} (\alpha^{-1} + \alpha^{-1})
\]
\[
= - \left\{ \delta^\nu + \varepsilon^\nu + K_{\gamma}^{\delta} \right\}^{3/4+\nu-\gamma} \delta^{1/2}
\]
(77)
where $K_{\gamma}^{\delta} \overset{def}{=} 2 (L T)^{1/4} (\alpha^{-1} + \alpha^{-1})$. Since $Q_{\delta,\varepsilon}^{\delta,\varepsilon} \leq 0$, we also have that
\[
\tilde{Q}_{\delta,\varepsilon}^{\delta,\varepsilon} \leq 2^{1+\nu-\gamma} (L T^{\delta/\varepsilon})^{1/4} (\alpha^{-1} + \alpha^{-1});
\]
thus
\[
|\tilde{Q}_{\delta,\varepsilon}^{\delta,\varepsilon}| \leq \delta^\nu + \varepsilon^\nu + K_{\gamma}^{\delta} \delta^{1/2}.
\]
To proceed, we calculate that
\[
|X_t| = \exp \left[ \mathbb{R}(A_{\delta,\varepsilon} \tilde{Q}_{\delta,\varepsilon}^t) \right] = \exp \left[ \mathbb{R}(A_{\delta,\varepsilon} \tilde{Q}_{\delta,\varepsilon}^t) \right]
\]
where we have used the fact that $\tilde{Q}_{\delta,\varepsilon}^{\delta,\varepsilon}$ is real-valued. From (76) (and using the fact that $\theta \neq 0$), we have that there is an $A^* \in D(\nu, \gamma)$ such that if $(\delta, \varepsilon) \in A \in D(\nu, \gamma)$ where $A \geq A^*$, then
\[
|\tilde{G}_{\delta,\varepsilon}| \leq \frac{1}{2} \left( \frac{\sigma^2 \chi_{\varepsilon}}{2 \varepsilon^3} \right) \theta^2
\]
(recall that $\lim_{\varepsilon \downarrow 0} \chi_{\varepsilon} = 1$). Thus if $(\delta, \varepsilon) \in A \in D(\nu, \gamma)$ where $A \geq A^*$,
\[
\mathbb{R}(A_{\delta,\varepsilon}) \leq \frac{1}{\gamma^2} \left( \frac{\sigma^2 \chi_{\varepsilon}}{4 \varepsilon^3 \theta^2} \right) \theta^2.
\]
Thus $\lim_{\nu, \gamma} L > 0$ for all $L > 0$ (recall (69)).

Let’s bound $\mathcal{E}^{(2)}(\delta, \varepsilon, L)$. We have that

$$\mathcal{E}^{(2)}(\delta, \varepsilon, L) = \mathbb{E} \left[ \int_{s=0}^{\epsilon \bar{\chi}} Y_s X_s ds \right].$$

From (73),

$$|A_{\delta, \varepsilon}| \leq \frac{1}{\delta \varepsilon} \left[ m_1 + \frac{\varepsilon^2}{d_{1+\nu, \gamma}} f \right],$$

$$|a_{\delta, \varepsilon}| \leq \frac{1}{2} \sigma^2 \tilde{X}_s^2 \varepsilon^2 \exp \left[ \frac{\varepsilon^4}{\delta^4} \right] \frac{2 \tilde{X}_s^2 \varepsilon^2}{d_{1+\nu, \gamma}} f$$

$$+ K_{B, A} \exp \left[ \frac{\varepsilon^4}{\delta^4} \right].$$

Thus (recall (69))

$$\lim_{(\delta, \varepsilon) \in A \subset D(\nu, \gamma)} \mathcal{E}^{(2)}(\delta, \varepsilon, L) = 0$$

for every $L > 0$. 
Let’s finally bound \( \mathcal{E}^{(1)}(\delta, \varepsilon, L) \). If \( \rho^{\delta, \varepsilon, L} \geq \bar{L}^{\delta} \) and \( \tau^{\delta, \varepsilon} \leq \bar{L}^{\delta} \), then \( \bar{\tau}^{\delta, \varepsilon} = \tau^{\delta, \varepsilon} \). Hence (note that \( |\exp[\theta \xi^{\delta, \varepsilon}]| = 1 \))

\[
|\mathcal{E}^{(1)}(\delta, \varepsilon, L)| \\
\leq \{1 + C_1(\delta, \varepsilon, L)\} \\
\times \left\{\mathbb{P}(\rho^{\delta, \varepsilon, L} < \bar{L}^{\delta} \delta^2) + \mathbb{P}(\tau^{\delta, \varepsilon} > \bar{L}^{\delta} \delta^2)\right\} \\
+ \mathbb{E}\left[|A_{\bar{\varepsilon}, \varepsilon, \bar{Q}^{\delta, \varepsilon}} - 1| \chi_{\{\bar{\tau}^{\delta, \varepsilon} = \tau^{\delta, \varepsilon}\}}\right].
\]

Using the rescaling of Remark 4.1, we have that

\[
\mathbb{P}\left\{\rho^{\delta, \varepsilon, L} < \bar{L}^{\delta} \delta^2\right\} \\
= \mathbb{P}\left\{\sup_{0 \leq t \leq L^{\delta^2}/\varepsilon} |\lambda_1^{\delta, \varepsilon}| \geq (L^{\delta^2}/\varepsilon)^{1/4}\right\} \\
= \mathbb{P}\left\{\sup_{0 \leq t \leq L^{\delta^2}/\varepsilon} |\lambda_1^{\delta, \varepsilon}| \geq (L^{\delta^2}/\varepsilon)^{1/4}\right\}. 
\]

Since

\[
\lim_{(\delta, \varepsilon) \in \Lambda} \lim_{\delta^{\gamma(\gamma+1)/2} \to \infty} \frac{\delta^2}{\varepsilon} \geq \frac{\lim_{(\delta, \varepsilon) \in \Lambda} \lim_{\delta^{\gamma(\gamma+1)/2} \to \infty} \left(\frac{\delta}{\varepsilon}\right)^2}{\infty}, 
\]

Lemma D.2 implies that

\[
\lim_{(\delta, \varepsilon) \in \Lambda} \mathbb{P}\left\{\rho^{\delta, \varepsilon, L} < \bar{L}^{\delta} \delta^2\right\} = 0
\]

for all \( L > 0 \). We also have that if \( L > 1 \),

\[
\mathbb{P}\left\{\bar{\tau}^{\delta, \varepsilon} > \bar{L}^{\delta} \delta^2\right\} \\
\leq \mathbb{P}\left\{||\xi^{\delta, \varepsilon}| > (L - 1)\delta^2\right\} \\
\leq \mathbb{P}\left\{|\xi^{\delta, \varepsilon}| > (L - 1)\delta^2\right\} \\
\leq \frac{e^\delta}{\delta^{1+\nu-\gamma/\varepsilon}} \mathbb{E}\left[|\xi^{\delta, \varepsilon}|\right] \\
\leq \frac{\varepsilon^{\gamma/\delta}}{(L - 1)^{\delta}},
\]

so (using Lemma B.1),

\[
\lim_{L \to \infty} \lim_{(\delta, \varepsilon) \in \Lambda} \mathbb{P}\left\{\bar{\tau}^{\delta, \varepsilon} > \bar{L}^{\delta} \delta^2\right\} = 0.
\]

If \( \bar{\tau}^{\delta, \varepsilon} = \tau^{\delta, \varepsilon} \), then

\[
|A_{\bar{\varepsilon}, \varepsilon, \bar{Q}^{\delta, \varepsilon}}| \\
\leq \frac{1}{\delta^{\varepsilon}} \left|m_1 + \frac{\varepsilon^{\gamma}}{\delta^{1+\nu-\gamma/\varepsilon}} f\right| \\
\times \left\{\left(\varepsilon^{\gamma/\delta}\right) \delta^{\gamma/2} - (\varepsilon^{\gamma/\delta})\right\} \\
\leq \left|m_1 + \frac{\varepsilon^{\gamma}}{\delta^{1+\nu-\gamma/\varepsilon}} f\right| \\
\times \left\{\left(\frac{\delta^{\gamma/2}}{\varepsilon^{\gamma/\delta}}\right) \delta^{\gamma/2 - 1} + \left(\frac{\varepsilon^{\gamma/\delta}}{\varepsilon^{\gamma/\delta}}\right)^{1/2}\right\}
\]

where

\[
C_3(\delta, \varepsilon, L) \equiv \left|m_1 + \frac{\varepsilon^{\gamma}}{\delta^{1+\nu-\gamma/\varepsilon}} f\right| \\
\times \left\{\left(\frac{\delta^{\gamma/2}}{\varepsilon^{\gamma/\delta}}\right) \delta^{\gamma/2 - 1} + \frac{\varepsilon^{(\gamma+1)/2}}{\delta}\right\}^{1/2}
\]

It is easy to see that

\[
\lim_{(\delta, \varepsilon) \in \Lambda} \mathbb{P}\left\{\rho^{\delta, \varepsilon, L} < \bar{L}^{\delta} \delta^2\right\} = 0 \\
\text{for each } L > 0, \text{ so collecting things together, we get that}
\]

\[
\lim_{L \to \infty} \lim_{(\delta, \varepsilon) \in \Lambda} \mathbb{P}\left\{|\mathcal{E}^{(1)}(\delta, \varepsilon, L)| = 0\right\}. 
\]

The claim now follows.

\[\square\]

\section{Proofs for epochs}

We here give the proofs for Section 7.

We start with general convergence result.

\begin{proposition}
For any \( f \in C_b(\mathbb{R}^2) \),

\[
\lim_{(\delta, \varepsilon) \in \Lambda} \mathbb{E}\left[f\left(\frac{\kappa_B(\delta, \varepsilon) \xi_B \delta^{1+\nu-\gamma/\varepsilon}}{\kappa_0(\delta, \varepsilon) \delta^{1+\nu-\gamma/\varepsilon}} \xi_A \delta^{1+\nu-\gamma/\varepsilon}\right)\right] \\
= \mathbb{E}\left[f(p_B V_B G_B, p_A V_A G_A)\right]. 
\]
\end{proposition}
Proof. It suffices to show that for all \( \theta_B \) and \( \theta_A \) in \( \mathbb{R} \),

\[
\lim_{(\delta, \varepsilon) \in A} \mathbb{E} \left[ \exp \left[ \frac{\theta_B}{\varepsilon} \frac{\kappa_{B, \delta, \varepsilon}}{\delta} \xi_{B, \delta, \varepsilon} + \frac{\theta_A}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right] \right]
\]

\[
= \exp \left[ -\frac{1}{2} \theta_B^2 \mu_A^2 V_{B}^2 \right] \exp \left[ -\frac{1}{2} \theta_A^2 \mu_A^2 V_{A}^2 \right] .
\]

This follows fairly quickly from Proposition 5.1. Indeed, under the assumption that \( \xi_{A, \delta, \varepsilon} \) and \( \xi_{B, \delta, \varepsilon} \) are independent, we have that

\[
\mathbb{E} \left[ \exp \left( \frac{\theta_B}{\varepsilon} \frac{\kappa_{B, \delta, \varepsilon}}{\delta} \xi_{B, \delta, \varepsilon} + \frac{\theta_A}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right) \right]
\]

\[
- \exp \left[ -\frac{1}{2} \theta_B^2 \mu_A^2 V_{B}^2 \right] \exp \left[ -\frac{1}{2} \theta_A^2 \mu_A^2 V_{A}^2 \right] \leq \mathcal{E}_{A, \delta, \varepsilon} + \mathcal{E}_{B, \delta, \varepsilon}
\]

where

\[
\mathcal{E}_{A, \delta, \varepsilon} \overset{\text{def}}{=} \mathbb{E} \left[ \exp \left[ \frac{\theta_B}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right] \right]
\]

\[
- \exp \left[ -\frac{1}{2} \theta_B^2 \mu_A^2 V_{B}^2 \right] \leq \mathbb{E} \left[ \exp \left[ \frac{\theta_B}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right] \right]
\]

\[
- \mathbb{E} \left[ \exp \left[ \frac{\theta_B}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right] \right] - \mathbb{E} \left[ \exp \left[ \frac{\theta_B}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right] \right]
\]

\[
- \exp \left[ -\frac{1}{2} \theta_B^2 \mu_A^2 V_{B}^2 \right] \leq |\theta_A| \frac{\kappa_{A, \delta, \varepsilon}}{\delta} - \mathbb{E} \left[ |\xi_{A, \delta, \varepsilon}| \right]
\]

\[
+ \mathbb{E} \left[ \exp \left[ \frac{\theta_B}{\varepsilon} \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon} \right] \right] - \exp \left[ -\frac{1}{2} \theta_B^2 \mu_A^2 V_{B}^2 \right]
\]

We use Proposition 5.1 on the second term and Lemma B.1 on the first. Collect things together to get the claim. \(\square\)

We now have the proofs of Corollaries 7.2 and 7.3.

**Proof of Corollary 7.2.** Define

\[
\Xi_{\delta, \varepsilon} \overset{\text{def}}{=} \frac{\kappa_{B, \delta, \varepsilon}}{\delta} \xi_{B, \delta, \varepsilon} - \frac{\kappa_{A, \delta, \varepsilon}}{\delta} \xi_{A, \delta, \varepsilon};
\]

then

\[
\{ \iota_{\delta, \varepsilon} \leq \iota_{B, \delta, \varepsilon} \} = \{ \Xi_{\delta, \varepsilon} \leq 0 \}
\]

and

\[
\{ \iota_{\delta, \varepsilon} \geq \iota_{A, \delta, \varepsilon} \} = \{ \Xi_{\delta, \varepsilon} \geq 0 \}.
\]

Since \((x, y) \mapsto x - y\) is a continuous map from \(\mathbb{R}^2\) to \(\mathbb{R}\), Proposition C.1 implies the weak limit

\[
\lim_{(\delta, \varepsilon) \in A} \Xi_{\delta, \varepsilon} = p_B V_B \mathcal{G}_B - p_A V_A \mathcal{G}_A.
\]

Since \(\mathbb{P}\{p_B V_B \mathcal{G}_B - p_A V_A \mathcal{G}_A = 0\}\) and

\[
\mathbb{P}\{p_B V_B \mathcal{G}_B - p_A V_A \mathcal{G}_A \leq 0\} = \mathbb{P}\{p_B V_B \mathcal{G}_B - p_A V_A \mathcal{G}_A \geq 0\} = \frac{1}{2},
\]

the claim follows. \(\square\)

**Proof of Corollary 7.3.** Since the map \((x, y) \mapsto \min\{x, y\}\) is a continuous map from \(\mathbb{R}^2\) to \(\mathbb{R}\), we have the claim. \(\square\)

We next identify the asymptotic mean of \(\hat{\mu}_{\delta, \varepsilon}\).

**Proof of Lemma 7.5.** Let’s first check that \(m_{\delta, \varepsilon}\) converges to \(m\). For each \(\ell > 0\), define \(\varphi_\ell \in C_0(\mathbb{R})\) as

\[
\varphi_\ell(z) \overset{\text{def}}{=} \begin{cases} 
\ell & \text{if } z > \ell \\
0 & \text{if } -\ell \leq z \leq \ell \\
-\ell & \text{if } z < -\ell.
\end{cases}
\]

We have that

\[
\sum_{z \in \mathbb{R}} \hat{\mu}(dz) - \sum_{z \in \mathbb{R}} \hat{\mu}_{\delta, \varepsilon}(dz)
\]

\[
= \left\{ \int_{z \in \mathbb{R}} \varphi_\ell(z) \hat{\mu}(dz) - \int_{z \in \mathbb{R}} \varphi_\ell(z) \hat{\mu}_{\delta, \varepsilon}(dz) \right\}
\]

\[
+ \left\{ \int_{z \in \mathbb{R}} \varphi_\ell(z) \hat{\mu}_{\delta, \varepsilon}(dz) - \int_{z \in \mathbb{R}} \varphi_\ell(z) \hat{\mu}_{\delta, \varepsilon}(dz) \right\}
\]

\[
+ \left\{ \int_{z \in \mathbb{R}} \varphi_\ell(z) \hat{\mu}_{\delta, \varepsilon}(dz) - \int_{z \in \mathbb{R}} \varphi_\ell(z) \hat{\mu}_{\delta, \varepsilon}(dz) \right\}.
\]
By dominated convergence,
\[ \lim_{\ell \to \infty} \int_{z \in \mathbb{R}} \varphi_\ell(z) \mu(dz) = \int_{z \in \mathbb{R}} z \mu(dz) \]
and by weak convergence,
\[ \lim_{(\delta, \epsilon) \in A} \left\{ \int_{z \in \mathbb{R}} \varphi_\ell(z) \mu(dz) - \int_{z \in \mathbb{R}} \varphi_\ell(z) \mu_{\delta, \epsilon}(dz) \right\} = 0. \]
We also have that
\[ \left| \int_{z \in \mathbb{R}} \varphi_\ell(z) \mu_{\delta, \epsilon}(dz) - \int_{z \in \mathbb{R}} z \mu_{\delta, \epsilon}(dz) \right| \leq \int_{|z| > \ell} (|z| - \ell) \mu_{\delta, \epsilon}(dz) \]
\[ \leq \int_{|z| > \ell} |z| \mu_{\delta, \epsilon}(dz) \leq \frac{1}{\ell} \int_{z \in \mathbb{R}} |z|^2 \mu_{\delta, \epsilon}(dz), \]
so that
\[ \lim_{\ell \to \infty} \lim_{(\delta, \epsilon) \in A} \left| \int_{z \in \mathbb{R}} \varphi_\ell(z) \mu_{\delta, \epsilon}(dz) - \int_{z \in \mathbb{R}} z \mu_{\delta, \epsilon}(dz) \right| = 0. \]
First take $\delta$ and $\epsilon$ to zero such that $(\delta, \epsilon) \in A \in \mathcal{D}(P^*)$ and then take $\ell \to \infty$; we get that indeed
\[ \lim_{(\delta, \epsilon) \in A} m_{\delta, \epsilon} = m. \]
Secondly, let’s evaluate $m$. For $c$ and $d$ in $(0, \infty)$, define
\[ I(c, d) \overset{\text{def}}{=} \frac{1}{2\pi} \int_{x = -\infty}^{\infty} \int_{y = -\infty}^{\infty} \frac{e^{-x^2/2}e^{-y^2/2} dy dx}{(x^2 + (dy))^2}. \]
We compute that
\[ I(c, d) = \frac{d}{2\pi} \int_{x = -\infty}^{\infty} e^{-x^2/2} \frac{dy}{y^2 + (c/d)^2} \]
\[ + \frac{c}{2\pi} \int_{y = -\infty}^{\infty} e^{-y^2/2} \frac{dx}{x^2 + (d/c)^2} \]
\[ = \frac{d}{2\pi} \int_{x = -\infty}^{\infty} \exp \left[ -\frac{x^2}{2} \left( 1 + (c/d)^2 \right) \right] dx \]
\[ - \frac{c}{2\pi} \int_{y = -\infty}^{\infty} \exp \left[ -\frac{y^2}{2} \left( 1 + (d/c)^2 \right) \right] dy \]
\[ = - \frac{d}{\sqrt{2\pi(1 + (c/d)^2)}} - \frac{c}{\sqrt{2\pi(1 + (d/c)^2)}} \]
\[ = - \frac{c^2 + d^2}{\sqrt{2\pi}}. \]
Thus if $P^* \in \mathcal{P} \setminus \{(0, 1), (1, 0)\}$,
\[ \int_{x \in \mathbb{R}} x \mu(dx) = -I(p_B V_B, p_A V_A) = -\frac{p_B^2 V_B^2 + p_A^2 V_A^2}{2\pi}. \]
If $P^* \in \{(0, 1), (1, 0)\}$, we use the calculation
\[ \int_{x = -\infty}^{0} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \frac{1}{\sqrt{2\pi}} \int_{x = -\infty}^{0} xe^{-x^2/2} dx \]
\[ = - \frac{1}{\sqrt{2\pi}}. \]
\[ \square \]

We finally start the proof of Theorem 7.6. We proceed by rewriting things a bit. For $t > 0$,
\[ Z^\delta_\epsilon - \hat{Z}_t = \frac{1}{\bar{\kappa}(\delta, \epsilon)/\delta^2} \left\{ \langle M^\delta_\epsilon \rangle_t - \frac{t}{T} \right\} + \frac{mt}{T^2} \]
\[ = U^{a, \delta, \epsilon}_t + U^{b, \delta, \epsilon}_t \]
where
\[ U^{a, \delta, \epsilon}_t \overset{\text{def}}{=} \frac{1}{\bar{\kappa}(\delta, \epsilon)/\delta^2} \left\{ \langle M^\delta_\epsilon \rangle_t - \frac{t}{T} \right\} \]
\[ \in \left\{ \frac{1}{T} \left( 1 - \frac{\bar{\kappa}(\delta, \epsilon)}{\delta^2} \right) \right\} \]
\[ = \frac{t}{T} \left( 1 - \frac{\bar{\kappa}(\delta, \epsilon)}{\delta^2} \right) \]
\[ \times \left\{ \hat{T} - \langle \bar{\kappa}(\delta, \epsilon)/\delta^2 \rangle m_{\delta, \epsilon} \right\} \]
\[ \times \left\{ (m - m_{\delta, \epsilon}) + \frac{\bar{\kappa}(\delta, \epsilon)}{\delta^2} m_{\delta, \epsilon} \right\} \]
\[ \times \left\{ \frac{m - m_{\delta, \epsilon}}{\delta^2} \right\} \]}
Clearly
\[
\lim_{(\delta, \varepsilon) \in A} \sup_{A \in \mathcal{D}(P^*)} \left| U^\delta_{t} \right| = 0. \tag{78}
\]

Let’s next replace absolute time by trading time.

**Lemma C.2.** We have that
\[
\lim_{(\delta, \varepsilon) \in A} \mathbb{P} \left\{ \tau_{2\delta^{-2}(T/\T)} \cap [T, \T] \right\} = 0.
\]

**Proof.** If \( \tau_{2\delta^{-2}(T/\T)} \cap [T, \T] \), then
\[
\left| \sum_{n=1}^{2\delta^{-2}(T/\T)} \left( \delta^2 \T + \bar{\varepsilon}(\delta, \varepsilon) \xi_n \right) \right| < T;
\]
in this case
\[
T > \left| [2\delta^{-2}(T/\T)] \delta^2 \T + \bar{\varepsilon}(\delta, \varepsilon) \right|
\]
\[
\times \sum_{n=1}^{2\delta^{-2}(T/\T)} \xi_n \xi_n > 2T + \bar{\varepsilon}(\delta, \varepsilon)
\]
\[
\times \sum_{n=1}^{2\delta^{-2}(T/\T)} \xi_n \xi_n.
\]

Hence by Markov’s inequality,
\[
\mathbb{P} \left\{ \tau_{2\delta^{-2}(T/\T)} \cap [T, \T] \right\} \leq \mathbb{P} \left\{ \tau_{2\delta^{-2}(T/\T)} \right\} \left( \sum_{n=1}^{2\delta^{-2}(T/\T)} \xi_n \right) > T
\]
\[
\leq \frac{\bar{\varepsilon}(\delta, \varepsilon)}{T} \left| \sum_{n=1}^{2\delta^{-2}(T/\T)} \xi_n \xi_n \right|
\]
\[
= \frac{\bar{\varepsilon}(\delta, \varepsilon)}{T} \left| \sum_{n=1}^{2\delta^{-2}(T/\T)} \xi_n \xi_n \right|
\]
\[
\leq \frac{\bar{\varepsilon}(\delta, \varepsilon)}{T} \left| \sum_{n=1}^{2\delta^{-2}(T/\T)} \xi_n \xi_n \right| \sqrt{K A.7.4}
\]
for \((\delta, \varepsilon) \in A \in \mathcal{D}(P^*)\) (use Lemma 7.4). The claim follows.

Let’s also bound the regularity of \( U^\alpha_{t} \) over about \( \delta^{-2} \) trades.

**Lemma C.3.** For each \( \kappa > 0 \),
\[
\lim_{(\delta, \varepsilon) \in A} \mathbb{P} \left\{ \max_{1 \leq n \leq [2\delta^{-2}(T/\T)]} \sup_{t_1 \leq t \leq t_2 = \delta^{-2} \T} \left| U^\alpha_{t_1} - U^\alpha_{t_2} \right| \geq \kappa \right\} = 0.
\]

**Proof.** Assume that \( \tau_{\delta^{-2}(T/\T)} \leq t \leq \tau_{\delta^{-2}(T/\T)} \) for some \( 1 \leq n \leq [2\delta^{-2}(T/\T)] \). Then
\[
\left| U^\alpha_{t_1} - U^\alpha_{t_2} \right| = \frac{1}{(\bar{\varepsilon}(\delta, \varepsilon)/\delta^2)(T + (\bar{\varepsilon}(\delta, \varepsilon)/\delta^2)m_{\delta, \varepsilon})} \left| t - \tau_{\delta^{-2}(T/\T)} \right|
\]
\[
\leq \frac{\delta^2}{\bar{\varepsilon}(\delta, \varepsilon) T + (\bar{\varepsilon}(\delta, \varepsilon)/\delta^2)m_{\delta, \varepsilon}}
\]
\[
\leq \frac{\delta^2}{\bar{\varepsilon}(\delta, \varepsilon) T + (\bar{\varepsilon}(\delta, \varepsilon)/\delta^2)m_{\delta, \varepsilon}} \max_{1 \leq n \leq [2\delta^{-2}(T/\T)]} \left| \xi_n \right|.
\]

Assume that \( \delta^4/\bar{\varepsilon}(\delta, \varepsilon) < \kappa/2 \) and \( (\bar{\varepsilon}(\delta, \varepsilon)/\delta^2)m_{\delta, \varepsilon} > 0 \).

Then
\[
\mathbb{P} \left\{ \max_{1 \leq n \leq [2\delta^{-2}(T/\T)]} \sup_{t_1 \leq t \leq t_2 = \delta^{-2} \T} \left| U^\alpha_{t_1} - U^\alpha_{t_2} \right| \geq \kappa \right\}
\]
\[
\leq \mathbb{P} \left\{ \max_{1 \leq n \leq [2\delta^{-2}(T/\T)]} \left| \xi_n \right| \right\}
\]
\[
> \frac{\kappa}{2\delta^2} \left| \bar{\varepsilon}(\delta, \varepsilon)/\delta^2 \right|m_{\delta, \varepsilon}
\]
\[
\leq \mathbb{P} \left\{ \bigcup_{n=1}^{[2\delta^{-2}(T/\T)]} \left| \xi_n \right| \right\}
\]
\[
> \frac{\kappa}{2\delta^2} \left| \bar{\varepsilon}(\delta, \varepsilon)/\delta^2 \right|m_{\delta, \varepsilon}
\]
Lemma C.4. We have that

$$\lim_{(\delta, \epsilon) \in \mathcal{A}} \mathbb{E} \left[ \max_{0 \leq n \leq T} \left| \frac{U_{\delta, \epsilon}^*}{\tau_n^\delta \epsilon} \right| \right] = 0.$$

Proof. For each $n \in \mathbb{N}$, define

$$u_n^{\delta, \epsilon} \overset{\text{def}}{=} U_{\delta, \epsilon}^* \left( \frac{1}{\delta^2} \frac{n \delta^2}{T + (\tilde{\xi}(\delta, \epsilon)/\delta^2)m_{\delta, \epsilon}} \right) = \frac{1}{\tilde{\xi}(\delta, \epsilon)/\delta^2} \left\{ n \delta^2 - \sum_{n' = 1}^{\delta^2 T} \left[ \frac{\delta^2 \tilde{\xi}(\delta, \epsilon)/\delta^2}{T + (\tilde{\xi}(\delta, \epsilon)/\delta^2)m_{\delta, \epsilon}} \right] \right\}$$

and set $\varphi_n^{\delta, \epsilon} \overset{\text{def}}{=} \left\{ \begin{array}{ll} \delta^2 \tilde{\xi}(\delta, \epsilon)/\delta^2 & \text{for } 1 \leq n' \leq n, \\ \delta^2 \tilde{\xi}(\delta, \epsilon)/\delta^2 & \text{for } n' > n. \end{array} \right.$ We have that

$$\mathbb{E}[u_{n+1}^{\delta, \epsilon} | \varphi_n^{\delta, \epsilon}] = u_n - \frac{\delta^2}{T + (\tilde{\xi}(\delta, \epsilon)/\delta^2)m_{\delta, \epsilon}} \mathbb{E} \left[ \tilde{\xi}_{n+1}^{\delta, \epsilon} - m_{\delta, \epsilon} | \varphi_n^{\delta, \epsilon} \right] = 0.$$

Thus $u_n^{\delta, \epsilon}$ is a martingale with respect to $\varphi_n^{\delta, \epsilon}$. By standard calculations,

$$\mathbb{E} \left[ \max_{0 \leq n \leq T} \left| \frac{U_{\delta, \epsilon}^*}{\tau_n^\delta \epsilon} \right| \right]$$

$$\leq \frac{\delta^2}{T + (\tilde{\xi}(\delta, \epsilon)/\delta^2)m_{\delta, \epsilon}} \left\{ \frac{\delta^2}{T + (\tilde{\xi}(\delta, \epsilon)/\delta^2)m_{\delta, \epsilon}} \mathbb{E} \left[ \tilde{\xi}_{n+1}^{\delta, \epsilon} - m_{\delta, \epsilon} | \varphi_n^{\delta, \epsilon} \right] \right\} = 0.$$

Finally, we bound $\mathbb{E}[u_n^{\delta, \epsilon} | \varphi_n^{\delta, \epsilon}]$ at the trade times.

$\square$
D. Stability of CIR processes

We here want to understand some stability issues of CIR processes. We want to show that the \( \lambda^{\pm, \varepsilon} \)’s of (18) are sufficiently bounded (thus giving a corresponding bound on the original \( \lambda^{\pm, \varepsilon} \)’s).

**Lemma D.1.** For each \( p > 0 \), there is a \( K_{p, D.1} \) such that

\[
E \left[ |\lambda^{\pm, \varepsilon}_t|^p \right] \leq K_{p, D.1}
\]

and

\[
E \left[ \sup_{0 \leq s \leq t} |\lambda^{\pm, \varepsilon}_s|^p \right] \leq K_{p, D.1} \left( 1 + \frac{t}{\varepsilon} \right)^{p/2}
\]

for all \( t \geq 0 \) and \( \varepsilon > 0 \).

We also want a tighter bound if we look only at the probability that \( \lambda^{\pm, \varepsilon} \) is large.

**Lemma D.2.** For each \( c > 0 \),

\[
\lim_{T \to \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} |\lambda^{\pm, \varepsilon}_t|^c \geq T^c \right\} = 0.
\]

To proceed, fix \( \alpha, \sigma, \lambda, \) and \( \lambda_0 \) in \( (0, \infty) \), and let \( \lambda \) satisfy the SDE

\[
d\lambda_t = -\alpha (\lambda_t - \lambda) \, dt + \sigma \sqrt{\lambda_t} \, dW_t \quad t > 0
\]

\[
\lambda_0 = \lambda_c
\]

where \( W \) is a standard Brownian motion; by Remark 4.1 the law of \( \lambda^{\pm, \varepsilon} \) can be described by speeding up a process like \( \lambda^* \).

First of all, define \( Z_t^{\lambda} \overset{\text{def}}{=} \lambda_t - \lambda \) and rewrite things a bit. We have that

\[
Z_t = (\lambda_0 - \lambda) - \alpha \int_{s=0}^{t} Z_s \, ds + \sigma \int_{s=0}^{t} \sqrt{\lambda_s} \, dW_s.
\]

This implies that

\[
Z_t = (\lambda_0 - \lambda) e^{-\alpha t} + \sigma \int_{s=0}^{t} e^{-\alpha (t-s)} \sqrt{\lambda_s} \, dW_s
\]

which in turn means that

\[
\lambda_t = Z_t + \lambda = \lambda_0 e^{-\alpha t} + \lambda (1 - e^{-\alpha t}) + \sigma \int_{s=0}^{t} e^{-\alpha (t-s)} \sqrt{\lambda_s} \, dW_s.
\]

(80)

Since \( \lambda_t \geq 0 \), we can get an easy bound on \( E[|\lambda_t|] \) by taking the expectation of (80). We get that

\[
E[|\lambda_t|] = E[\lambda_t] = \lambda_0 e^{-\alpha t} + \lambda (1 - e^{-\alpha t}) \leq \lambda_0 + \lambda.
\]

(81)

Let’s proceed by showing that all moments of \( \lambda \) are uniformly bounded in time.

**Lemma D.3.** For each \( p > 0 \),

\[
\sup_{t \geq 0} E[|\lambda_t|^p] < \infty.
\]

Proof. By (81), we definitely have the claim if \( p = 1 \). Next note that from (79),

\[
Z_t^{2p} = (\lambda_0 - \lambda)^{2p} - 2p\alpha \int_{s=0}^{t} Z_s^{2p} \, ds + 2p \sigma \int_{s=0}^{t} Z_s^{2p-1} \sqrt{\lambda_s} \, dW_s
\]

\[
+ \sigma^2 \bar{p} (2\bar{p} - 1) \int_{s=0}^{t} Z_s^{2p-2} \lambda_s \, ds;
\]

for all \( \bar{p} \geq 1 \); thus

\[
E[Z_t^{2p}] = (\lambda_0 - \lambda)^{2p} - 2p\alpha \int_{s=0}^{t} E[Z_s^{2p}] \, ds + \sigma^2 \bar{p} (2\bar{p} - 1) \int_{s=0}^{t} E[Z_s^{2p-2} \lambda_s] \, ds.
\]

(82)
Let’s first take \( p = 1 \). We get that
\[
\mathbb{E}[Z^p_t] = (\lambda_0 - \bar{\lambda})^2 - 2\alpha \int_0^t \mathbb{E}[Z^p_s] ds + \sigma^2 \int_0^t \mathbb{E}[\lambda_0] ds.
\]
Solving this and using (81), we hence have that
\[
\mathbb{E}[Z^2_t] = (\lambda_0 - \bar{\lambda})^2 e^{-2\alpha t} + \sigma^2 \int_0^t e^{-2\alpha (t-s)} \mathbb{E}[\lambda_0] ds
\leq (\lambda_0 - \bar{\lambda})^2 + \frac{\sigma^2}{2\alpha} (\lambda_0 + \bar{\lambda})
\]
and consequently
\[
\mathbb{E}[\lambda^2_t] \leq 2 \{ \mathbb{E}[Z^2_t] + \bar{\lambda}^2 \}
\leq 2 \left( (\lambda_0 - \bar{\lambda})^2 + \frac{\sigma^2}{2\alpha} (\lambda_0 + \bar{\lambda}) + \bar{\lambda}^2 \right).
\]
Thus the claim holds if \( p = 2 \).

Fix now \( p \geq 2 \) and assume that \( K \defeq \sup_{t \geq 0} \mathbb{E}[|\lambda_t|^{2p-1}] \) is finite (which we just proved if \( p = 2 \)). We want to use (82). From Young’s inequality, we have that
\[
\sigma^2 (2p - 1) \mathbb{E}[Z^{2p}_s - 2\lambda_s] 
\leq \alpha \mathbb{E}[Z^{2p}_s] + C_1 \mathbb{E}[\lambda^p_s]
\leq \alpha \mathbb{E}[Z^{2p}_s] + C_1 \mathbb{E}[|\lambda_s|^{2(2p-1)/2}]
\leq \alpha \mathbb{E}[Z^{2p}_s] + C_1 K^{p/(2p-1)}
\]
where
\[
C_1 \defeq (\frac{\sigma^2 (2p - 1)}{p}) \left( \frac{p - 1}{p\alpha} \right)^{-1}.
\]
Thus
\[
\mathbb{E}[Z^{2p}_s] \leq (\lambda_0 - \bar{\lambda})^{2p} - p\alpha \int_0^t \mathbb{E}[Z^{2p}_s] ds + C_1 K^{p/(2p-1)}
\]
and hence
\[
\mathbb{E}[Z^{2p}_t] \leq (\lambda_0 - \bar{\lambda})^{2p} e^{-p\alpha t}
+ \int_0^t e^{-p\alpha (t-s)} C_1 K^{p/(2p-1)} ds
\leq (\lambda_0 - \bar{\lambda})^{2p} + \frac{C_1}{p\alpha} K^{p/(2p-1)}.
\]
Consequently
\[
\mathbb{E}[|\lambda_t|^{2p}]
\leq 2^{p-1} \left\{ \mathbb{E}[Z^{2p}_t] + \bar{\lambda}^{2p} \right\}
\leq 2^{p-1} \left\{ (\lambda_0 - \bar{\lambda})^{2p} + \frac{C_1}{p\alpha} K^{p/(2p-1)} + \bar{\lambda}^{2p} \right\}.
\]
By induction, we get the claim for \( p \) even, and by Jensen’s inequality, we then get the claim for all \( p \).

Let’s now consider running suprema of the moments of \( \lambda \) (see the statement of Lemma D.1).

**Lemma D.4.** For each \( p > 0 \), there is a \( K_p > 0 \) such that
\[
\mathbb{E} \left[ \sup_{0 \leq s \leq t} |\lambda_s|^p \right] \leq K_p (1 + t^{p/2}).
\]

**Proof.** Define the martingale
\[
M_t \defeq \int_0^t \sqrt{\lambda_s} dW_s;
\]
then we can rewrite (80) as
\[
\lambda_t = \lambda_0 e^{-\alpha t} + \bar{\lambda} (1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha (t-s)} dM_s. \tag{83}
\]
Integrating the stochastic integral by parts,
\[
\int_0^t e^{-\alpha (t-s)} dM_s = e^{-\alpha t} \int_0^t e^{\alpha s} dM_s
= e^{-\alpha t} \left\{ e^{\alpha t} M_t - \alpha \int_0^t e^{\alpha s} M_s ds \right\}
= M_t - \alpha \int_0^t e^{-\alpha (t-s)} M_s ds.
\]
Define \( M^*_t \defeq \sup_{0 \leq s \leq t} |M_s| \) for all \( t \geq 0 \); then
\[
\left| \int_0^t e^{-\alpha (t-s)} dM_s \right| \leq M^*_t + \alpha \int_0^t e^{-\alpha (t-s)} M^*_s ds \leq 2M^*_t.
\]
Thus for any $t \geq 0$, (83) implies that
\[
\sup_{0 \leq s \leq t} |\lambda_s| \leq \lambda_0 + \bar{\lambda} + 2\sigma M_t^*.
\]

Thus for $\bar{p} \geq 1$,
\[
E \left[ \sup_{0 \leq s \leq t} |\lambda_s|^{2\bar{p}} \right] \leq 3^{2\bar{p}-1} \left\{ \lambda_0^{2\bar{p}} + \bar{\lambda}^{2\bar{p}} + 2^{2\bar{p}}\sigma^{2\bar{p}} E[|M_t^*|^{2\bar{p}}] \right\}
\]
\[
\leq 3^{2\bar{p}-1} \left\{ \lambda_0^{2\bar{p}} + \bar{\lambda}^{2\bar{p}} + K \sigma^{2\bar{p}} \left( \int_{s=0}^t \lambda_s ds \right)^\bar{p} \right\}
\]
\[
\leq 3^{2\bar{p}-1} \left\{ \lambda_0^{2\bar{p}} + \bar{\lambda}^{2\bar{p}} + K \sigma^{2\bar{p}} \left( \int_{s=0}^t E[|\lambda_s|^\bar{p}] ds \right) \right\}
\]
\[
\leq 3^{2\bar{p}-1} \left\{ \lambda_0^{2\bar{p}} + \bar{\lambda}^{2\bar{p}} + K \sigma^{2\bar{p}} \sup_{s \geq 0} E[|\lambda_s|^\bar{p}] \right\}
\]

where $K$ is a constant from the Burkholder-Davis-Gundy inequalities (Karatzas & Shreve, 1991, Chapter 3.3). Using Lemma D.3, we get the claimed result for even $\bar{p}$; use Jensen’s inequality to get the result for the remaining $\bar{p}$.

Proof of Lemma D.1. Collect things together, rescaling time as in Remark 4.1.

Let’s next bound the probability that $\lambda^\pm$ exceeds a large level before too long a time (i.e., let’s prove Lemma D.2). To proceed, consider a general CIR process
\[
d\zeta_t = -\left( \zeta_t - (1 + \theta) \right) dt + \sqrt{2\kappa} dW_t \quad t \geq 0
\]
\[
\zeta_0 = \zeta_0.
\] (84)

Any CIR process can be rescaled (via rescaling time and space) to get (84) for some $\bar{\theta} \in \mathbb{R}$. We want to bound the probability that $\zeta$ exceeds a large level before too long a time. We assume that $\bar{\theta} > 0$; the Feller conditions then imply that $\zeta$ is strictly positive (and by the rescaling of Lemma D.2, this corresponds to (15)). For $L > 0$, define
\[
\tau_L \overset{\text{def}}{=} \inf \{ t \geq 0 : \zeta_t \geq L \}.
\]

For any $T > 0$,
\[
P\{ \tau_L \leq T \} = P \left\{ 1 - \frac{\tau_L}{T} \geq 0 \right\} \leq e^{\tau_L/T} E \left[ e^{-\tau_L/T} \chi_{(\tau_L < \infty)} \right].
\]

Note that the generator of (84) is the operator
\[
(Lf)(x) = x f'(x) - (x - (1 + \bar{\theta})) f(x) \quad x > 0
\]
for $f \in C^2(0, \infty)$.

For $p > 0$ and $q > -1$, define
\[
I_{p,q}(x) \overset{\text{def}}{=} \int_{t=0}^1 e^{xt} t^{p-1} (1-t)^q dt \quad x \geq 0
\]

Note that $I_{p,q} \in C^\infty(0, \infty)$ with
\[
\dot{I}_{p,q}(x) = I_{p+1,q}(x). \quad x > 0
\] (85)

If $p$ and $q$ are positive we can integrate by parts and see that
\[
x I_{p+1,q}(x)
\]
\[
= \int_{t=0}^1 x e^{xt} t^p (1-t)^q dt
\]
\[
= \int_{t=0}^1 e^{xt} \left\{ -pt^{p-1} (1-t)^q + qt^p (1-t)^{q-1} \right\} dt
\]
\[
= -p I_{p,q}(x) + q I_{p+1,q-1}(x) \quad \text{for all } x > 0.
\] (86)

Finally,
\[
I_{p,q}(x) - I_{p+1,q}(x) = \int_{t=0}^1 e^{xt} t^{p-1} (1-t)^{q+1} dt
\]
\[
= I_{p,q+1}(x) \quad x > 0
\] (87)

for all $p > 0$ and $q > -1$.

Let’s now apply $L$. Fix $\kappa \in (0, \bar{\theta})$. Using (85), (86) and (87), we have that
\[
(L_{\kappa, \theta})(x)
\]
\[
= x I_{\kappa+1, \theta-\kappa}(x) - x I_{\kappa+1, \theta-\kappa}(x) + (1 + \bar{\theta}) I_{\kappa+1, \theta-\kappa}(x)
\]
\[
= \left\{ -\kappa I_{\kappa+1, \theta-\kappa}(x) + (\bar{\theta} - \kappa) I_{\kappa+1, \theta-\kappa}(x) \right\}
\]
\[
= \left\{ -\kappa I_{\kappa+1, \theta-\kappa}(x) + (\bar{\theta} - \kappa) I_{\kappa+1, \theta-\kappa}(x) \right\}
\]
\[
= \kappa I_{\kappa+1, \theta-\kappa}(x) + (\bar{\theta} - \kappa) I_{\kappa+1, \theta-\kappa}(x)
\]
\[
= \kappa I_{\kappa+1, \theta-\kappa}(x) - (\bar{\theta} - \kappa)
\]
\[
\left\{ I_{\kappa+1, \theta-\kappa}(x) - I_{\kappa+1, \theta-\kappa}(x) - I_{\kappa+1, \theta-\kappa}(x) \right\}
\]
\[
= \kappa I_{\kappa+1, \theta-\kappa}(x).
\] (88)
for all \( x > 0 \); i.e., \( I_{\kappa,0,\kappa} \) solves a Laguerre differential equation. Assume now that \( T > 1/\theta \) and set

\[
\Phi_{T,L}(x) \overset{\text{def}}{=} \frac{L_{1/T,\theta_1-1/T}(x)}{L_{1/T,\theta_1-1/T}(L)}. \quad x > 0
\]

Comparing this with (88), we have that

\[
(L\Phi_{T,L})(x) = \frac{1}{T} \Phi_{T,L}(x) \quad x > 0
\]

\[
\Phi_{T,L}(L) = 1.
\]

Returning to (84), the process \( \{\Phi_{T,L}(\zeta_t)e^{-t/T}; t \geq 0\} \) is a martingale so

\[
\Phi_{T,L}(\zeta_0) = \mathbb{E} \left[ \Phi_{T,L}(\zeta_{t\wedge \tau_L})e^{-(t\wedge \tau_L)/T} \right].
\]

By bounded convergence,

\[
\Phi_{T,L}(\zeta_0) = \mathbb{E} \left[ \Phi_{T,L}(\zeta_{t\wedge \tau_L})e^{-(t\wedge \tau_L)/T} \chi_{t\wedge \tau_L < \infty} \right]
\]

and consequently

\[
\mathbb{P}\{\tau_L \leq T\} \leq e\Phi_{T,L}(\zeta_0). \quad \tag{89}
\]

To understand \( \mathbb{P}\{\tau_L \leq T\} \) for large \( T \) and \( L \) we thus need to understand the asymptotics of \( I \).

Let’s first get a simple upper bound on \( I \). For \( t \in [0,1], e^{-t} \leq e^x \) and \((1-t)^p \leq 1\). Thus

\[
I_{p,q}(x) \leq e^x \int_{t=0}^{1} t^{p-1} dt = \frac{e^x}{p}.
\]

To bound \( I_{p,q} \) from below, let’s first make the change of variables \( u = t^p \); then

\[
I_{p,q}(x) = \frac{1}{p} \int_{u=0}^{1} \exp [F_{x,p,q}(u)] du
\]

where

\[
F_{x,p,q}(u) \overset{\text{def}}{=} xu^{1/p} + q \ln(1 - u^{1/p})
\]

for \( u \in (0, 1) \). We calculate that

\[
F'_{x,p,q}(u) = \frac{1}{p} u^{1/p-1} \left\{ x - \frac{q}{1 - u^{1/p}} \right\}
\]

\[
F''_{x,p,q}(u) = \frac{1}{p} \left( \frac{1}{p} - 1 \right) u^{1/p-2} \left\{ x - \frac{q}{1 - u^{1/p}} \right\} - \frac{u^{2(1-p)/p}}{p^2} \left( \frac{q}{1 - u^{1/p}} \right)^2.
\]

Let’s now assume that \( x \geq q \), and define

\[
u^*_p(x,p,q) \overset{\text{def}}{=} \left( 1 - \frac{q}{x} \right)^p;
\]

then \( \nu^*_p(x,p,q) \in (0,1) \) and

\[
F'_{x,p,q}(\nu^*_p(x,p,q)) = 0
\]

and

\[
F''_{x,p,q}(\nu^*_p(x,p,q)) = -\frac{x^2}{pq^2} \left( 1 - \frac{q}{x} \right)^{2(1-p)} < 0;
\]

thus

\[
F_{x,p,q}(u) \leq F_{x,p,q}(\nu^*_p(x,p,q)) = x - q + q \ln \frac{q}{x} \quad u \in (0,1)
\]

If \( 0 \leq u \leq \nu^*_p(x,p,q) \), we have that

\[
\frac{1}{1 - u^{1/p}} \leq \frac{1}{1 - (\nu^*_p(x,p,q))^{1/p}} = \frac{x}{q}
\]

and hence if additionally \( p < 1 \),

\[
F''_{x,p,q}(u) \geq -q \frac{u^{2(1-p)/p}}{p^2(1 - u^{1/p})^2} \geq -\frac{q(u^*_p(x,p,q))^{2(1-p)/p}}{q^{p^2}}
\]

\[
= -\frac{(u^*_p(x,p,q))^{2(1-p)/p} x^2}{q^2}.
\]

Thus if \( 0 \leq u \leq \nu^*_p(x,p,q) \), we can write that

\[
F_{x,p,q}(\nu^*_p(x,p,q)) - F_{x,p,q}(u)
\]

\[
= - \int_{r=\nu^*_p(x,p,q)}^{u} F''_{x,p,q}(r)(\nu^*_p(x,p,q) - r) dr
\]

\[
\leq \frac{(u^*_p(x,p,q))^{2(1-p)/p} x^2}{2q} (u^*_p(x,p,q) - u)^2
\]

so

\[
F_{x,p,q}(u) \geq x - q + q \ln \frac{q}{x} - \frac{(u^*_p(x,p,q))^{2(1-p)/p} x^2}{2q} (u^*_p(x,p,q) - u)^2.
\]
Thus

\[ I_{p,q}(x) \geq \exp \left[ x - q + q \ln \frac{2}{p} \right] \int_{u=0}^{\frac{\sigma_{p,q}^2}{p}} \frac{(u_{x,p,q})^{2(1-p)}}{2q} \exp \left[ \frac{\left( \frac{u_{x,p,q}^2}{2q} \right)}{2} \right] du \]

\[ = e^{-q^2} \int_{x=-\frac{\sigma_{p,q}^2}{p}}^{0} \exp \left[ \frac{-\left( \frac{u_{x,p,q}^2}{2q} \right)}{2} \right] dz. \]

Let’s now prove something like Lemma D.2.

**Lemma D.5.** For any \( c' > 0 \), we have that

\[ \lim_{T \to \infty} \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \zeta_t \geq T^{c'} \right\} = 0. \]

**Proof.** Using (89), we have (if \( T > 1/\bar{\theta} \)) that

\[ \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \zeta_t \geq T^{c'} \right\} \leq e^{\Phi_{1/T} \bar{\theta} - 1/T} \left( e^{T^{c'}} (\bar{\theta} - 1/T)^{\bar{\theta} - 1/T} \right) \leq T^{c'} \bar{c} \left( e^{T^{c'}} (\bar{\theta} - 1/T)^{\bar{\theta} - 1/T} \right) \leq \left\{ \begin{array}{l} e^{T^{c'}} (\bar{\theta} - 1/T)^{\bar{\theta} - 1/T} \\ \times \left( \int_{z=-u_{x,p,q}^2/2q}^{0} \exp \left[ \frac{-\left( \frac{u_{x,p,q}^2}{2q} \right)}{2} \right] dz \right) \end{array} \right\}^{-1} \]

\[ = \frac{e^{\bar{\theta} - 1/T} \bar{c}}{(\bar{\theta} - 1/T)^{\bar{\theta} - 1/T}} e^{c} e^{-c} T^{c} (\bar{\theta} - 1/T + 1)^{\bar{\theta} - 1/T + 1} \times \left\{ \begin{array}{l} \int_{z=-u_{x,p,q}^2/2q}^{0} \exp \left[ \frac{-\left( \frac{u_{x,p,q}^2}{2q} \right)}{2} \right] dz \end{array} \right\}^{-1} \]

We note that

\[ \lim_{T \to \infty} \sup_{0 \leq t \leq T} \zeta_t = \lim_{T \to \infty} \left( 1 - \frac{\bar{\theta} - 1/T}{T^{c'}} \right)^{1/T} = 1 \]

and thus

\[ \lim_{T \to \infty} \frac{(\sup_{0 \leq t \leq T} \zeta_t)^{2(1-1/T)}}{2(\bar{\theta} - 1/T)} = \frac{1}{20} \]

\[ \lim_{T \to \infty} -u_{x,p,q}^2/2q \leq -\infty. \]

The claim now follows.

**Proof of Lemma D.2.** Set \( \bar{\theta} \triangleq \frac{2 \alpha_1 + \lambda_0}{\sigma_1^2} - 1 \) and \( \zeta_c = \frac{2 \alpha_1 + \lambda_0}{\sigma_1^2} \); then \( \{ \lambda_t^c ; t \geq 0 \} \) has the same law as \( \left\{ \frac{\sigma_2^2}{2 \alpha_1^2} \zeta_{\alpha_t^c} ; t \geq 0 \right\} \). Thus if \( c' \in (0, c) \),

\[ \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \lambda_t^c \geq T^{c'} \right\} = \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \zeta_{\alpha_t^c} \geq \frac{\alpha_{c'}^c}{20 T^{c'} \sigma_1^2} \right\} \leq \mathbb{P} \left\{ \sup_{0 \leq t \leq T} \zeta_{\alpha_t^c} \geq \left( \frac{\alpha_{c'}^c}{20 T^{c'} \sigma_1^2} \right)^{c'} \right\} \]

if \( T^{c'-c} \geq \frac{2 \alpha_{c'}^c}{20 \sigma_1^2 \sigma_1^2} \). The conclusion now follows from Lemma D.5.

**References**


