# CUSUM Statistics applied to Market Microstructure

Ivan E. Perez

January 18, 2021

## 1 Introduction

This proposal falls under the scope of changepoint detection applied to market microstructure. CUSUM algorithms using Brownian motion have been applied for trend following strategies[9, 2]. Stopping time problems using Levy processes have been studied by Rodosthenous[10] and Dayanik and Sezer[5]. Empirical studies [8, 16] have been the center of market microstructure research but and considerable effort has been dedicated to developing its underlying microeconomic theory OHara[11, 3].

Our work will operate at the intersection of these three fields. We will first compile theoretical work and empirical research to understand fundamental order book dynamics. We will then seek to develop observable random variables that arise from a change in price due to a private signal. We will develop a detection scheme to categorize periods exhibiting adverse changes in real time. With these constructs we will develop a tractable framework and to explain how market participants respond to private signals.

The proposal is organized as follows, In **Section 2** We introduce the limit order book and in **Section 2.1** we introduce market microstructure models that describes the order book as a Quasi Birth-Death (QBD) process. In **Section 2.2** We describe the Glosten market participant framework, and empirical findings from Cetin and Haelworeck[3]. In **Section 3** we outline our detection scheme and research plan.

## 2 Background

### 2.1 The limit order book and the Huang queue-reactive model

We describe a limit order book as defined in Huang[7]. The order book is on a price grid with $2K$ quotes. The bid side is denoted by a negative index $[Q_{-i} = 1, \ldots, K]$, and the ask is denoted as $[Q_i = 1, \ldots, K]$, where $Q_{\pm i}$ represents the limit at $+i$ ticks to the right, $-i$ ticks to the left of the mid price $p_{mid} = (Q_{-1} + Q_1)/2$. The volume in shares at each limit is given as $q_i$.

The order book queue sizes at the best 3 bids and asks is modeled as a QBD process, $X_t = \{q_{-K}, \ldots, q_{-1}, q_1, \ldots, q_K\}$. It is assumed that arrivals and departures at the first limit are independent from the second and third limits, and for non-empty queues at the first limit, the rate of departures is larger than the rate of arrivals. At each limit $i$, the rate of arrival is described using $f_i(q)$, and departures with $g_i(q)$. The initial rates are estimated using the maximum likelihood estimator for Poisson processes, and are modified using the function $\mathcal{S}_{m,l}(x; m, l)$ that takes constants $m$ and $l$ to be the $33^{\text{rd}}$ and $66^{\text{th}}$ percentile of the cumulative distribution of the historical queue size. The arrival and departure functions become:

$$f_i(q) = \lambda_i^L(q, \mathcal{S}_{m,l}(q)) \tag{1}$$

$$g_i(q) = \lambda_i^M(q, \mathcal{S}_{m,l}(q)) + \lambda_i^C(q, \mathcal{S}_{m,l}(q)). \tag{2}$$

Their results fall in line with their empirical invariant distribution of queue sizes at the first and second limits, allowing them to estimate the probability of execution of market and limit orders for a given state of the order book. However, their methodology of fitting an invariant distribution over a long period only represents a broad aggregation of investor behavior. To gain some insight into how specific market events (i.e., a private signal) causes three classes of investors to respond, we recount the Glosten market participant framework[6].

## 2.2 Glosten market participant framework

The Glosten framework[6], takes place in a single period and has three classes of investors. 1. the competitive liquidity supplier, 2. competitive dealer and 3. $N \geq 1$ informed traders who know liquidation value $V$ of the asset. The initial limit order book is generated by liquidity suppliers that places limit orders. $h : \mathbb{R} \to \mathbb{R}$ is the function faced by market order when hitting a limit order level $y$. In the Glosten framework the market maker already has $Z \sim N(0, \sigma^2)$ number of shares of the asset, assumed to be independent of V. The market maker's role of smoothly moving the price to the true liquidation value to $V$ is done through buying and selling shares from liquidity providers that create the initial order book.

Cetin[3] et. al show that the liquidation value with $N$ informed traders,

$$V = E^v \left[ \frac{h(Z + Nx^*)}{N} + \frac{N-1}{N^2 x^*} \int_0^{Nx^*} h(Z+u)du \right] \tag{3}$$

was determined from first order conditions where $x^*$ is the optimal demand for a single informed trader. $E^v$ denotes the expectation operator of the investor where $V = v$.

Section 5.2 Figure 9 of [3] shows that for many informed traders, $N \geq 5$, the bid ask spread increases with the arrival of unbounded (e.g., log-normal) signals. Despite the rationale that for a fixed $V$ an increasing $N$ would result in decreasing $x^*$, traders may have minimum order sizes, and are likely not be fully informed as to the number of other insiders that have received the private signal. Cetin et al. also states that market makers would change the market impact of ahead of a move to $V$ as a result of a private signal.

In the following section we propose experiments to verify the claim. **Section 2.1** has given us a broad understanding of how order flow can be described using a QBD process, and the results above explain how market participants behave in accordance to private signals. In the following section we develop a detection scheme to identify shifts to $V$ using CUSUM processes.

# 3 Research Plan

We propose an initial experiment to evaluate one of the observations in **Section 3**. Using Crobat [12] with ACoPrA[14] we can record instantaneous changes in order book depth on free exchanges. We would seek to determine if a period in which spreads increase and there is a marked move in price to $V$ would be preceded by an increase in order book depth. Similar to Dufrense [4] we take order book depth as a proxy for price impact.

Our detection scheme would implement a Radon-Nikodym derivative to detect a change of measure from $\lambda_0\mu \to \lambda_1\mu, \forall \lambda_1\mu > \lambda_0\mu$ for a Compound Poisson random variable $\lambda\mu$ as described in [15] for:

1. rate of order book depth increase(decrease),
2. effective spread, and
3. rate of market order arrival increase(decrease).

The rate of limit order arrivals(1), $\lambda$ in orders/second would be the rate of arrivals for limit order insertion net of cancellations on a single side (i.e., bid or ask) of the market. $\mu$ would be the mean size of net limit orders in an interval. The base distribution for the effective spread(2) would need to be further understood, but [1] implies that it may be Poisson-like. The rate of market order arrival(3) would be identical to the work I have previously done in [13], where the base rate is the rate of market order arrivals per second.

The detection scheme would declare periods where a marked shift to $V$ is observed on a combination of CUSUM alarms. The initial experiment would reveal how market participants act on private signals prior to price change. We will still need to understand what governs the threshold order book depth or spread for a change to $V$? How is private information disseminated to informed traders? How do market makers manage queue sizes to reduce the market impact from informed traders? This project would expand our understanding of the optimal stopping time problem that traders face when processing private signals.

# References


[1] Tim Bollerslev, Ian Domowitz, and Jianxin Wang. Order flow and the bid-ask spread: An empirical probability model of screen-based trading. *Journal of Economic Dynamics and Control*, 21(8):1471 – 1491, 1997. Computational financial modelling.

[2] M. Carlisle, O. Hadjiliadis, and I. Stamos. *Handbook of High Frequency Trading and Modeling in Finance*. John Wiley & Sons Inc., 2016.

[3] Umut Cetin and Henri Waelbroeck. Informed trading, limit order book and implementation shortfall: Equilibrium and asymptotics. *ArXiv*, 2020.

[4] Pierre Collin-Dufresne and Vyacheslav Fos. Do prices reveal the presence of informed trading? Working Paper 18452, National Bureau of Economic Research, October 2012.

[5] Savas Dayanik and Semih O. Sezer. Sequential testing of simple hypotheses about compound poisson processes. *Stochastic Processes and thei Applications*, 116:1892–1919, 2006.

[6] Lawrence R. Glosten. Is the electronic open limit order book inevitable. *Journal of Finance*, 49:1127–1161, 1994.

[7] Weibing Huang, Charles-Albert Lehalle, and Mathieu Rosenbaum. Simulating and analyzing order book data: The queue-reactive model. *Journal of the American Statistical Association*, 110(509):107–122, 2015.

[8] Marc Mezard Jean-Phillippe Bouchand and Marc Potters. Statistical properties of stock order books: empirical results and models. *Quantitative Finance*, 2:251–256, 2006.

[9] K. Lam and H.C. Yam. Cusum techniques for technical trading in financial markets. *Asia-Pacific Financial Markets*, 4:257–274, 1997.

[10] Hongzhong Zhang Neofytos Rodosthenous. When to sell an asset amid anxiety about drawdowns. *Mathematical Finance*, 30:1422–1460, 2020.

[11] Maureen O'Hara. *Market Microstructure Theory*. John Wiley & Sons Inc., 1998.

[12] Ivan E. Perez. crobat, 2020.

[13] Ivan E. Perez. *A Study of CUSUM Statistics on Bitcoin Transactions*. PhD thesis, Hunter College - City University of New York, 2020.

[14] Tom Podlaski. Copra, 2019.

[15] Vincent H. Poor and Olympia Hadjiliadis. *Quickest Detection*. Cambridge University Press, 2009.

[16] Arseniy Kukanov Rama Cont and Sasha Stoikov. The price impact of orderbook events. *Journal of Financial Econometrics*, 12:47–88, 2014.