



Università  
della  
Svizzera  
italiana

Faculty  
of  
Informatics

## Bachelor Thesis

August 26, 2020

# Comets

Real-Time Financial Market Visual Analytics

Jason Naldi

---

### *Abstract*

Financial markets are the place where traders buy and sell financial products, such as stocks, bonds, and others. What happens in financial markets is very often public, meaning that every trader knows almost all the actions of all the other traders. The market then is like a source that emits a stream of data of all the actions taken by the traders inside of it.

To make money in a market, a trader needs to buy from someone and sell to someone else at a higher price than she paid for when buying. If a trader sells at a lower price than she paid for when buying, then she loses money. The difficulty of trading comes from the fact that one cannot know at which price she will be able to sell in the future, so one is left to guess about whether she will make money or not. Comets attempts to help traders in guessing by providing them with a heatmap showing all the events that occurred in a market in the given timeframe at any given price. The more a price was of interest, the more colored it appears in the heatmap, so traders can see at a glance where the action is and might be in the near future.

---

### **Advisor:**

Prof. Michele Lanza

### **Co-advisor:**

Dr. Roberto Minelli

---

Approved by the advisor on Date:

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Definitions</b>	<b>2</b>
<b>3</b>	<b>Problem</b>	<b>3</b>
<b>4</b>	<b>Comets — A Web App for Real-Time Analysis of Financial Markets</b>	<b>3</b>
4.1	Reading Comets . . . . .	5
4.2	Trading with Comets . . . . .	5
4.3	User Interactions . . . . .	5
<b>5</b>	<b>Implementation Details</b>	<b>6</b>
5.1	In a Nutshell . . . . .	6
5.2	Data Representation . . . . .	6
5.3	Precision Tricks . . . . .	7
5.4	Price and Time Scales . . . . .	7
5.5	Performance Strategies . . . . .	7
<b>6</b>	<b>Conclusions</b>	<b>11</b>

# 1 Introduction

In the last century, the world wide web brought a lot of disruption to many fields. Finance was one of the more drastically affected: In the past, traders stood on exchange floors and tried to shout the loudest in order to get heard by other traders; nowadays almost all traders sit quietly in front of computer screens and click “buy/sell” buttons. Although technology has improved some conditions for traders (like allowing them to always know the best bid and ask), technology has also made some things worse. In the days of trading on exchange floors, all traders were able to see and (often) hear each other. Hence, if traders saw the representative of a big bank walk towards the desk of a company, they could hear what trading that person was doing and this sometimes helped them better understand the market activity they would later see. In the computer age, traders can only know how much of what has been bought and sold. There is no longer a way to know who is the person behind the trades, unless she tells it publicly. This cloak of anonymity allowed a new kind of manipulative traders to emerge: Equipped with extremely deep pockets of money, these traders were—and still are—able to control the markets by using questionable techniques. For example, if a manipulator wants to play a stock to the upside, she would carry out a scheme like the following: Scare traders into selling by generating a lot of artificial selling activity. Take the other side of scared sellers and buy from them. Once the desired quantity of stock has been bought, stop the artificial selling activity. Then, scare traders into buying by generating a lot of artificial buying activity. Take the other side of all scared buyers and sell to them. Once all the stock has been sold, stop the artificial buying activity. As the trading day comes to an end, the traders who fell for the tricks of the manipulator have technically donated their money to him.

How can traders survive today, in a world where market manipulation is so easy to carry out? The answer is that manipulators leave trails, just like comets flying through the sky. If one follows a trail, she is falling behind, rushing to chase an event whose meaning she cannot see. If one is able to see the comet however, she can try to ride it and make a profit off of those who are missing it, or at least avoid falling for the trap.

## 2 Definitions

The world of trading is quite technical. In this section a few definitions are provided for common terms used in this report.

**Market:** A place (often digital) where people can buy and sell units of a product.

**Trader:** An entity that trades in a market.

**Fill:** The event where the trader trades with another trader. A fill has a price and an amount of base units exchanged.

**Market order:** An order where the submitter specifies the quantity of base units to buy/sell. The quantity is guaranteed to be filled, but the trader has no control over the fill price (depends on best bid/ask).

**Limit order:** An order where the trader specifies the quantity and price of base units to buy/sell. The quantity is not guaranteed to be filled, but the trader has full control over the fill price (filling depends on whether the trader's order is taken by someone else by means of an opposing market order).

**Best bid:** Limit buy order whose offered price is the highest at a given time in the market (best price for market sell orders).

**Best ask:** Limit sell order whose offered price is the lowest at a given time in the market (best price for market buy orders).

**Tape:** Chronologically ascending historic collection of market orders.

**Book/orderbook:** Collection of unfilled limit orders grouped by price (a limit order of size 100 at price 20 and another limit order of size 200 at price 20 are represented in the book with a size of 300 at price 20).

**Tick:** The maximum level of granularity for bid and ask prices. A tick size of 0.5 indicates that the next closest prices to 10 are 9.5 and 10.5.

**BitMEX:** a private company that hosts multiple online markets for contracts based on cryptocurrencies.

**XBTUSD:** the product that is most frequently traded on BitMEX.



Figure 1. Price chart for XBTUSD on BitMEX, each candle shows one minute of trading activity — 25 May 2019, 18:00 to 20:30.

### 3 Problem

There are two types of traders: those who buy and sell because they need, like a farmer trading the wheat she produced; and those who buy and sell in hope of making a profit, or at least not losing money. Comets is tailored for the second type of traders: the speculators (referred to as simply “traders”).

To make money, a trader needs to buy a product from someone and sell it to someone else at a higher price than she paid for it. If the trader sells at a lower price than she paid for the product, then she loses money. This is trading in a nutshell: the dynamics are extremely simple. However, things are not so easy, because traders have no way to know whether in the future they will be able to actually sell at a higher price or have to settle for selling at a lower price than they paid for when buying. So, the activity of trading carries a risk and traders are left to guess about the future prices of what they are buying and selling now. In fact, trading is a game of guessing.

The most common source of data on which traders base their guesses is price charts, which are a historic summary of the best prices at which a trader bought or sold a product within a given period of time (see Figure 1). In price charts, each vertical bar is called a “candle.” Each candle represents a period of trading and shows four data points, again for the given period of trading: first price, highest price, lowest price and last price; where “price” is either the price at which someone bought or sold.

The problem with price charts is that they show only a tiny fraction of all market activity. Moreover, of the data they show a lot is still lost, because price charts are a summary of market activity. For example, in a minute there might be more than four trades but a candle can only show four data points. Moreover, there is no indication as to whether the first and last prices shown in a candle were prices at which someone bought or sold. This loss of data can be so drastic that sometimes price charts become impossible to decipher (see Figure 1 for some examples).

Is there a way to view all market data without having to omit/compress any of it? And does this extra data help in better understanding price movements?

### 4 Comets — A Web App for Real-Time Analysis of Financial Markets

How can traders look at the data coming from a market and better understand what is going on? The goal of Comets is to assist traders in approaching this problem. At a high level, Comets is a web application that displays in real-time the evolution of market and limit orders for a given product using a custom rendering engine built in WebGL.

Figure 2 is showing the same price and time range for the same market as in Figure 1 (see Figure 3 for an overlay of the two). Thanks to the visualisation offered by Comets, the trader is now able to better understand market activity:

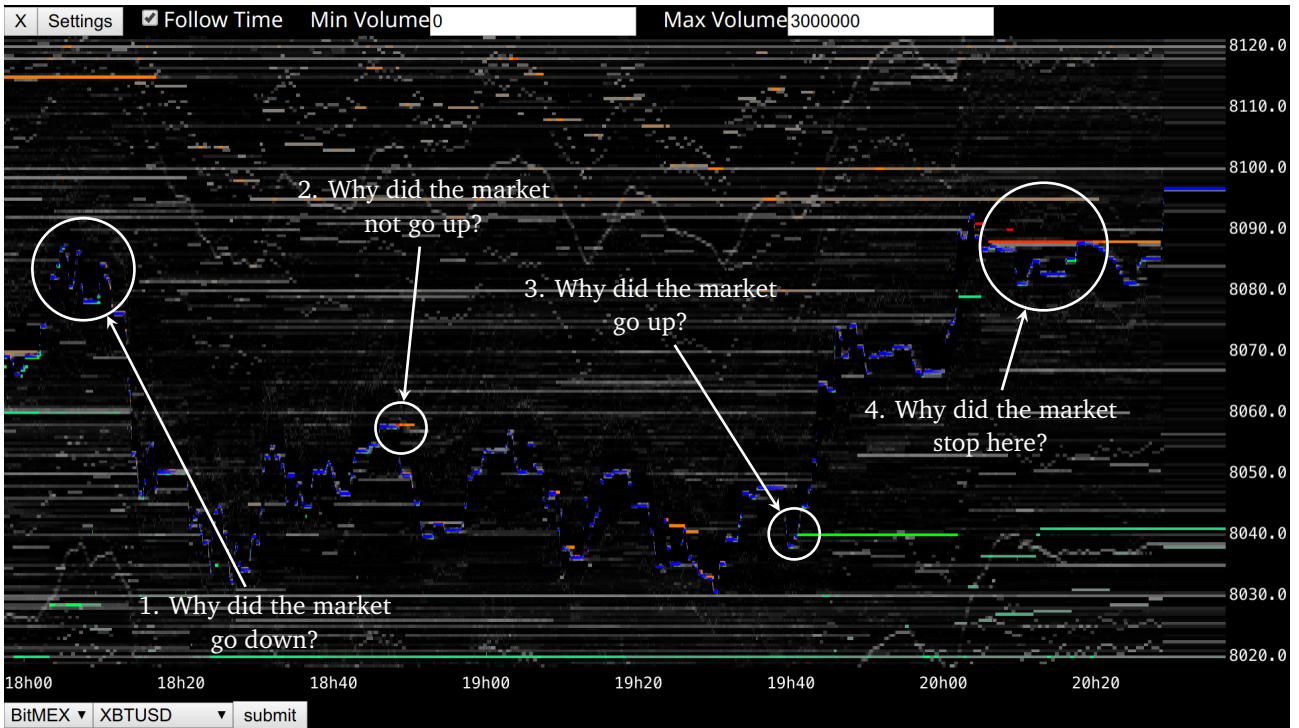


Figure 2. Comets displaying XBTUSD on BitMEX — 25 May 2019, from 18:00 to 20:30.

1. Why did the market go down? Because there were only a few big buyers that kept adding and removing their bids, until they stopped adding and simply disappeared. Then, new sellers came into the market.
2. Why did the market not go up? Because a big seller came into the market as buyers were raising their prices.
3. Why did the market go up? Because a very big buyer came into the market.
4. Why did the market stop here? Because a big seller came into the market to trade with the many buyers that started showing up.

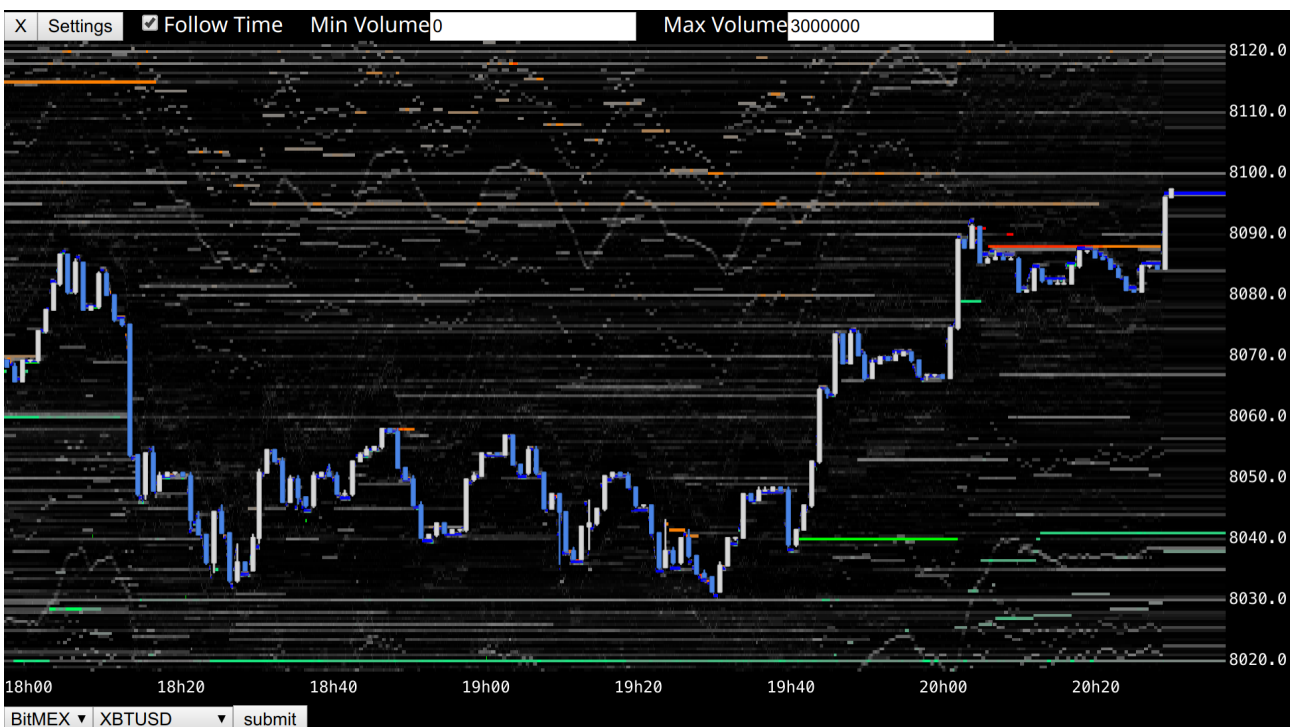


Figure 3. Price chart vs. Comets for XBTUSD on BitMEX — 25 May 2019, from 18:00 to 20:30.

## 4.1 Reading Comets

Comets is an application designed to help traders see the past and present state of financial markets. To do so, Comets shows all market events as colored bars rendered in an organised way: by time and by price at which events occurred. Each bar represents a price at which people can trade and whenever the color of the bar changes, it means that an event happened: if the color progressively becomes red, it means that many people want to sell at that price, whereas if the color progressively becomes green, it means that many people want to buy at that price. The blue bars represent the best price at which someone bought or sold at any point in time. These blue bars progressively become green if more and more people are buying at the same price, or they become red if more and more people are selling at the same price. Syntactically:

**Red line:** Large amount of limit sell orders.

**Green line:** Large amount of limit buy orders.

**Blue line:** Best price at which a trader bought or sold.

**Black/grey area:** Range of prices at which very few limit orders were present.

The color of each horizontal price bar changes over time because, as the market moves, traders change their opinions about its future and thus add or remove limit orders.

## 4.2 Trading with Comets

In Figure 4, notice the green line moving up (Figure 4.1): it is a large limit buy order and there were not enough sellers to fill it, so the owner raised its price multiple times.

Seeing this, a trader could have guessed that the order would have continued to be moved up and thus buy at 17:14, then sell at 17:17 after the buyer was filled and large sellers joined the market (Figure 4.3).

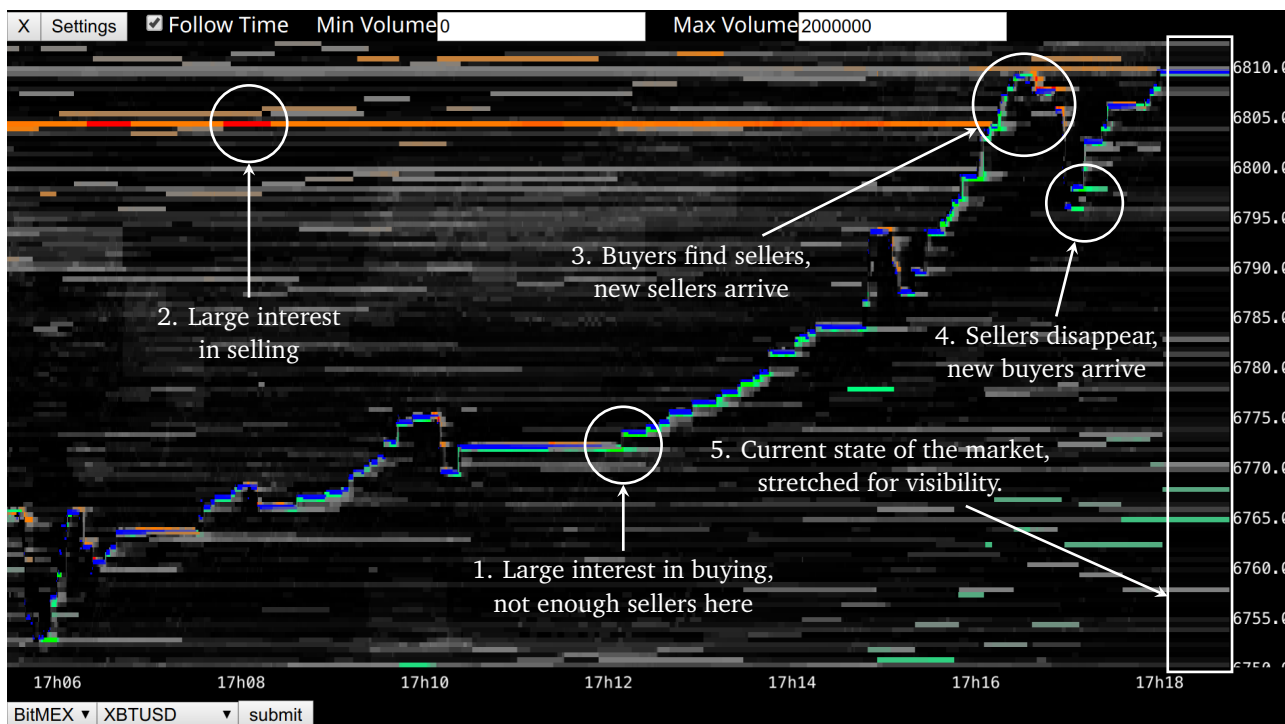


Figure 4. Comets displaying XBTUSD on BitMEX — 11 May 2019, from 17:06 to 17:18.

## 4.3 User Interactions

Comets is a web-based application designed to help traders better visualise what is happening inside of a financial market. To operate Comets, select the desired data provider and market, then click “submit.” The market view will appear, displaying the visible time and price ranges. To move around:

- j to move up by one tick

- k to move down by one tick
- J to move up by 10 ticks
- K to move down by 10 ticks
- u to move up by 50 ticks
- d to move down by 50 ticks
- + to increase the number of visible price levels
- - to decrease the number of visible price levels
- > to increase the number of visible time levels
- < to decrease the number of visible time levels

## 5 Implementation Details

### 5.1 In a Nutshell

Comets is a web application. All data is processed by some servers, whereas rendering is left to the client. The client is a web app fully written in JavaScript without relying on any external framework. To achieve better responsiveness, the size of all HTML elements is handled by flexboxes. Rendering is done in WebGL, because it is the only library which can handle the giant amount of data to be displayed. For example, 12 hours of trading activity can contain upwards of 12 million data points. For now, the back-end is a NodeJS application that relays market data to all connected clients.

### 5.2 Data Representation

Events in a market happen sequentially and atomically, meaning that whenever someone executes a limit or market order, no one else's order can be executed. This is necessary, because the entity hosting the market needs to keep track of exactly how much every trader owns, and failure to do so can result in either creating or destroying money. Given this condition, data coming from a market can be organised as a time series. In Comets, each of these time series is identified by:

- The market where data comes from
- The type of event that identifies the data (market buy, market sell, limit buy, etc.)
- The price at which the time series occurs

Each time series is an array of interleaved timestamps and values. Timestamps and values in each time series are represented as 64 bit floating point values, unless otherwise specified.

On a high level, data is stored as follows (where an arrow represents a key -> value relation and [data] represents an array of data):

```
market id -> event type -> price -> [data]
```

For example, given the following data:

```
6 3 8216 [ 1560492910900, 297503, 1560492910944, 296703 ]
```

the values represent:

- 6: id of the market—XBTUSD on BitMEX
- 3: type of event—limit sell
- 8216: price at which the event occurred
- 1560492910900, 297503: on June 14, 2019 at 6:15:10.900 GMT the amount of contracts to buy was 297,503
- 1560492910944, 296703: on June 14, 2019 at 6:15:10.944 GMT the amount of contracts to buy was 297,603.

Another example, given the following data:

```
6 2 8216 [ 1560492910903, 400, 1560492910941, 800 ]
```

the values represent:

- 6: id of the market—XBTUSD on BitMEX
- 3: type of event—limit sell
- 8216: price at which the event occurred
- 1560492910903, 400: on June 14, 2019 at 6:15:10.903 GMT a trader bought 400 contracts
- 1560492910941, 800: on June 14, 2019 at 6:15:10.941 GMT a trader bought 800 contracts.

### 5.3 Precision Tricks

When sent to the GPU, data is compressed into 32 bit floating point values due to limitations in GPU technology. To avoid losing significant precision, the rendered data is normalised on the fly. For example, if the current range of timestamps is  $[1,000,000,000; 1,000,020,000]$ , it will be compressed to  $[0; 20,000]$ . Compression is performed on the GPU, since this type of operation can be easily parallelised. To avoid loss of precision, 64 bit floating point values are converted into pairs of high and low bits, such that for a value  $v$ , where  $T = 2^{23}$  (highest value at which 32 bit floating points preserve unit precision):

- $\text{low} = v \bmod T$
- $\text{high} = v - \text{low}$

The 32 bit version of a 64 bit value  $v$  represented as  $v_{\text{high}}$  and  $v_{\text{low}}$ , given a center value  $C$  as  $C_{\text{high}}$  and  $C_{\text{low}}$  and  $T$ , is:  $(v_{\text{high}} - C_{\text{high}}) \cdot T + v_{\text{low}}$ .

The splitting of a value into its high and low bits is performed on the CPU, since 64 bit precision is lost when sending data to the GPU.

### 5.4 Price and Time Scales

Data by itself difficult to understand: There is a strong need to know what the data seen represents. To help in this task, a specific class (`ScaledDisplay`) takes care of rendering ranges of values with configurable formatting rules. These values for now are either prices or timestamps (formatted as user-readable dates).

The challenge behind rendering ranges of data is that one would expect to see enough information to reference each level, but at the same time, the rendered labels must not overlap each other (otherwise they are unreadable). To achieve this, an algorithm takes the range of values to be rendered, estimates the size of each label and greedily tries to fit as many labels as possible in the given screen space. If there is not enough space to render each label, the algorithm performs a compression operation where labels are displayed with intervals (for example, instead of rendering 1, 1.5, 2, 2.5, the algorithm may render 1, 2). This algorithm is designed with the user in mind: The labels must be easy to compare, especially when compression occurs. To this aim, the algorithm tries to fit labels to the closest multiples of 1, 2 and 5, starting from a baseline which is a multiple of the compression (this means, if the range is  $[12, 20]$  and the compression is 5, the algorithm will render 15, 20). Visually, Figures 5 and 6 show how Comets changes the interval of displayed labels for prices to make sure that each remains readable, whereas Figures 7 and 8 show the same compression but for timestamps.

### 5.5 Performance Strategies

The main challenge for the front end is the very large size of the time series to be rendered. Broadly, each time series is rendered as a sequence of colored rectangles. Each rectangle is composed of two triangles that need to be filled. WebGL is able to determine whether a triangle is off-screen and thus avoid rendering it. Yet, the amount of data to be rendered is too large, and the mere determining what to render is too expensive if left to the GPU. Moreover, as the amount of triangles to be processed grows, there is a high risk of saturating the memory bandwidth in the GPU, thus damaging performance (from observations, the saturation on a Sapphire RX 580 Nitro+ 8GB seems to be around 24 million triangles, which is more or less 12 hours of trading during an active day). To face these issues, the application instructs the GPU to render only the exact range that should be visible on the screen. The discovery



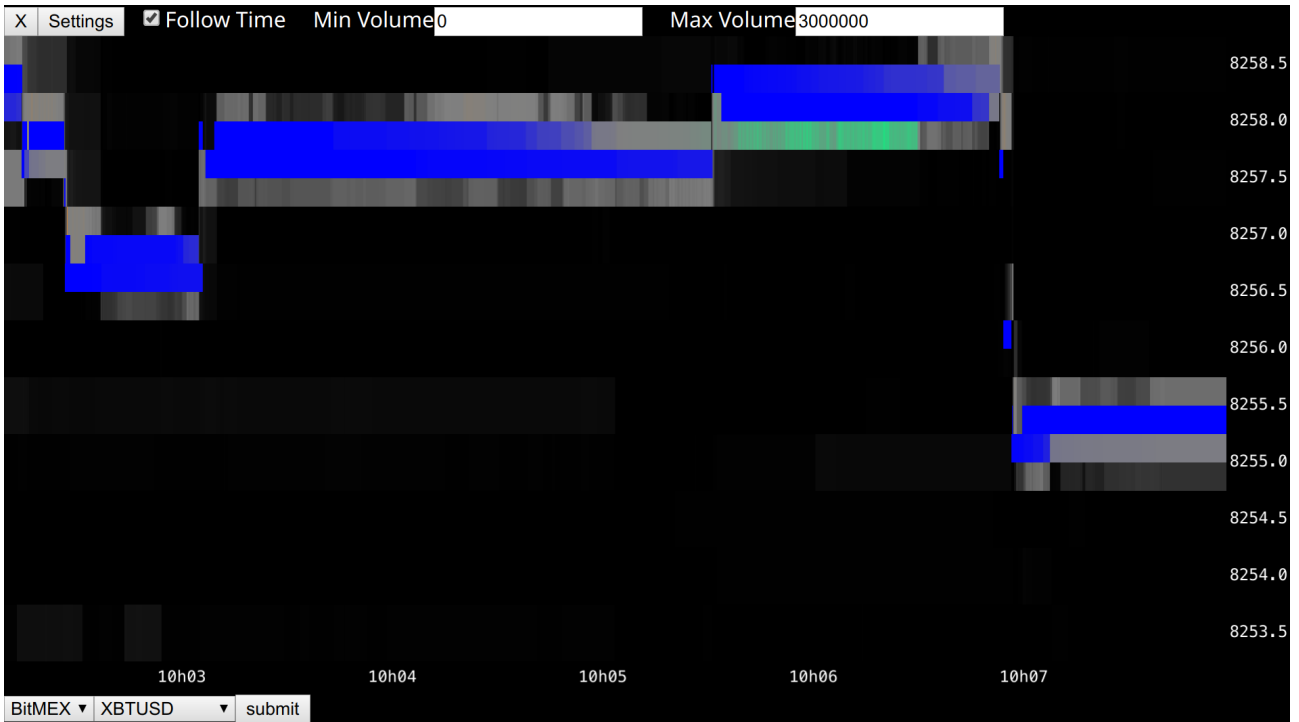


Figure 5. Comets displaying the label for each price bar.

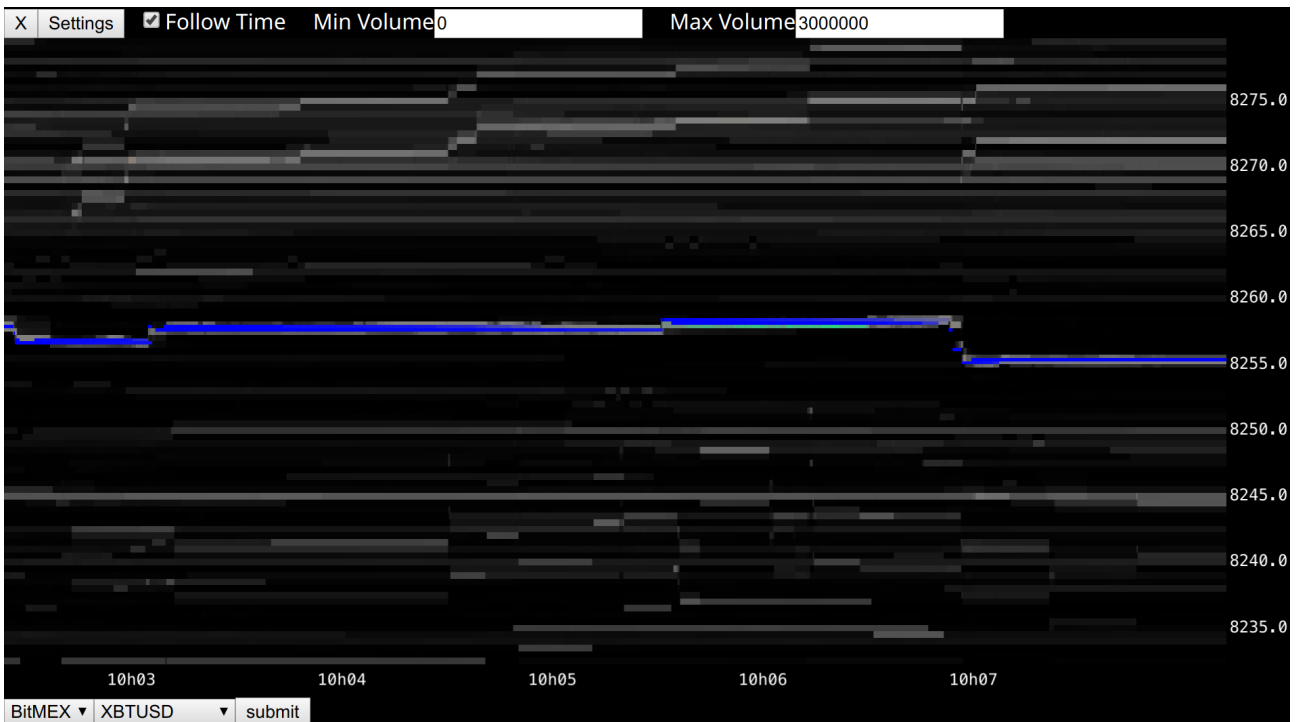


Figure 6. Comets displaying the label only for some price bars.

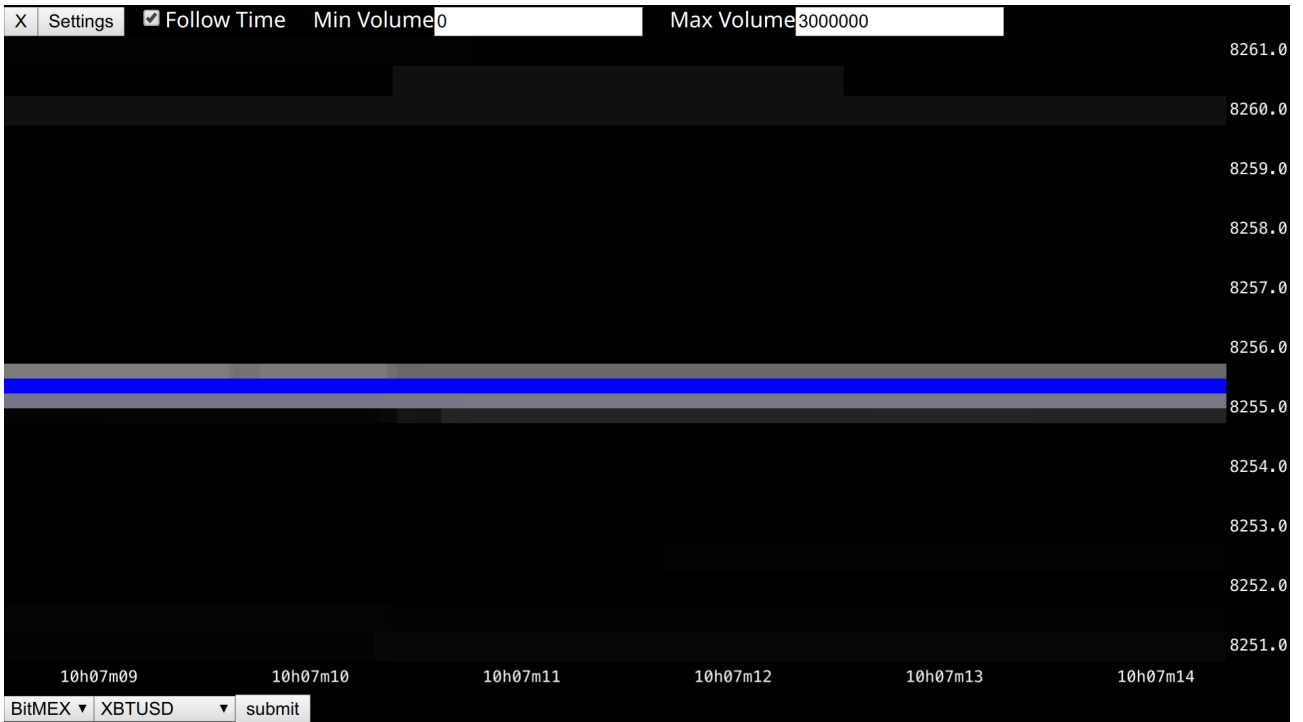


Figure 7. Comets displaying the label for timestamps by tens of milliseconds.

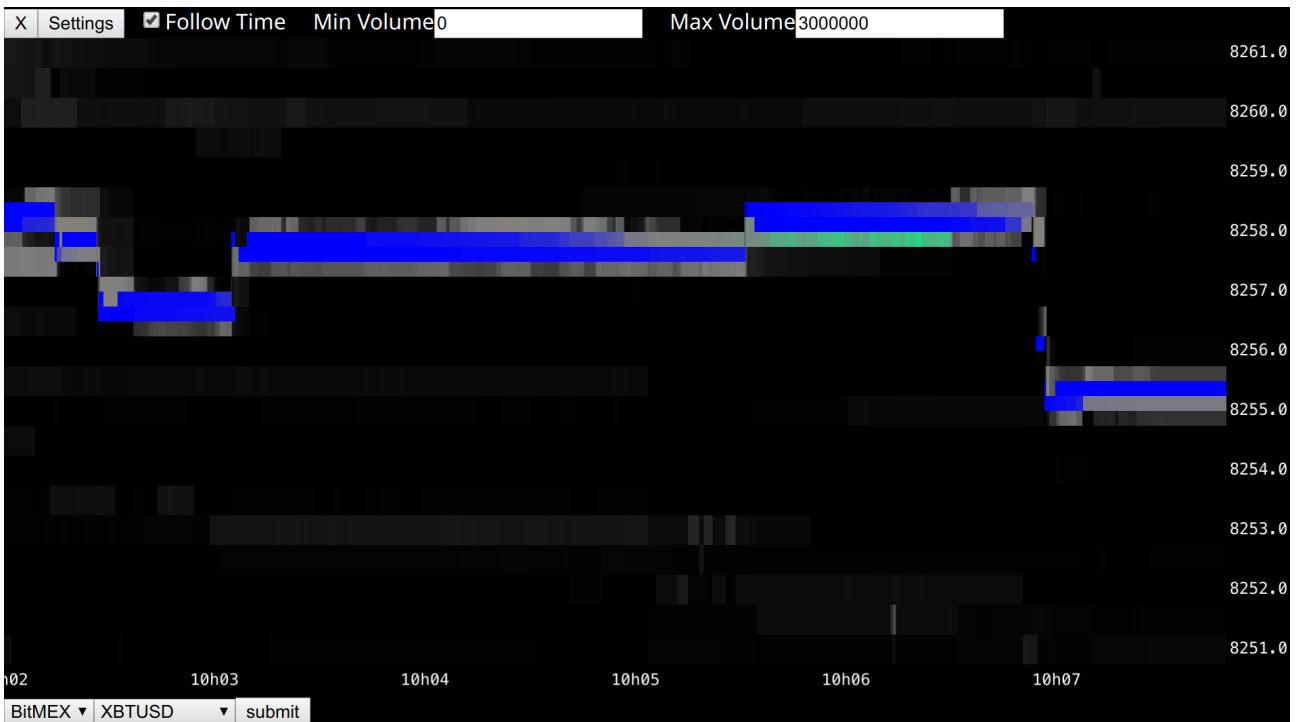


Figure 8. Comets displaying the label for timestamps by minutes.

of this range is not too expensive: time series are sorted by time, and finding a value can be done via binary search in  $O(\log(n))$  time [1].

Another technique for less-performant GPUs is that of caching what has already been rendered. This way, the number of triangles to render is greatly reduced, since the GPU has to only draw the new data between a frame and the previous and recycle the rendering of all historic data. This can be done only because historic data cannot change. This approach however has a few drawbacks: In some frames, pieces of data might not be visible, because they are too tiny to be fit in a pixel. When recycling rendered frames, data that was not visible is technically lost forever in the visualisation. Another problem is that—again due to pixels having a fixed, non-infinitely-small size—sometimes the rendered image is off by a pixel. There are further strategies to attempt and avoid these off-by-one errors, but they are not exact. To mitigate these problems, every few frames the application simply re-renders the whole screen. The result is that Comets is often able to sustain 60 frames per second even on low-performance GPUs such as the Intel Iris Pro Graphics 5200 (which comes in MacBook Pros from Mid 2015).

## 6 Conclusions

In today's financial markets, the speed of trading is extremely high, and so is the amount of data that can be generated for any given period of time. Traders (the people who buy and sell in financial markets hoping to make money) have to process all this data in very short periods of time and understand what is happening, so they can try and guess what will happen next. Comets attempts to help these traders by providing them with real-time visualisations of all the public events within a market. The result is an intuitive interface that allows traders to spot events very quickly, because all market data is shown as colored bars rather than numbers. Moreover, Comets shows all market data to its users, whereas existing alternatives (price charts) show only a tiny fraction of the data. This, in turn, can provide new trading opportunities that would not emerge when looking at price charts. It is important to note however that Comets is just a tool and, as such, it is up to the trader to make the best of it, just like a racing car requires a good pilot to win the race.

## References

- [1] D. Knuth. *The Art of Computer Programming, Volume 3: Sorting and Searching, Second Edition*. Addison-Wesley, 1998.