

# Market Making in Crypto

Sasha Stoikov, Elina Zhuang, Hudson Chen, Qirong Zhang, Shun Wang,  
Shilong Li, and Chengxi Shan

Cornell Financial Engineering Manhattan

December 20, 2024

## Abstract

We develop automated market-making algorithms for cryptocurrency perpetual contracts, which provide liquidity while managing risk and maximizing returns. Using historical candlestick data, we develop an alpha signal we call the Bar Portion (BP), which is robust across cryptocurrencies. We then use the Hummingbot<sup>1</sup> platform, an open-source framework for algorithm development, to fine tune risk management parameters before live trading. By live trading on the SOL-USDT, DOGE-USDT, and GALA-USDT trading pairs over a 24-hour period, we show that BP outperforms a baseline MACD signal.

## 1 Introduction

The exponential growth of cryptocurrency markets has heightened the need for efficient liquidity provision and risk management. Market makers are key participants, continually quoting buy and sell prices to maintain liquidity. This paper focuses on developing automated market-making algorithms for crypto perpetual contracts, aiming to improve liquidity and mitigate trading risks. The ultimate goal is to design strategies that optimize profitability while effectively controlling market risk.

At the heart of this project is the Hummingbot platform, an open-source solution for building, optimizing, backtesting, and deploying algorithmic trading strategies. Hummingbot is well-suited to the crypto ecosystem, where the combination of 24/7 trading and high volatility requires constant market engagement and real-time adaptability. By automating market making, Hummingbot enables traders to maintain liquidity and capitalize on short-term trading opportunities.

---

<sup>1</sup>We thank Federico Cardoso and Michael Feng at Hummingbot for introducing us to the platform and guiding us through the steps needed for live trading.

Market-making algorithms depend on several key parameters, including spread, order size, and the frequency of order placements, which must be carefully calibrated. Another vital component is high-frequency alpha generation, which seeks to exploit short-term price inefficiencies by executing trades rapidly with minimal latency.

Beyond alpha generation, this work also emphasizes risk management. A dedicated team worked on implementing a triple barrier strategy, a sophisticated approach to managing market risk. This strategy defines three price thresholds—take-profits, stop-losses, and time barriers—to systematically close positions based on predefined risk criteria. By combining alpha generation and the triple barrier approach, our work aims to develop algorithms that can thrive in the fast-paced, high-risk world of crypto trading.

## 2 Data Exploration

In this section, we explore the dataset used for developing and testing the market-making algorithms, focusing on the selection of the cryptocurrency universe, the characteristics of the data, and the initial empirical findings.

### 2.1 Universe Selection

We selected 30 cryptocurrencies based on their market capitalization, ensuring that only coins with perpetual contracts were included in our analysis. The rationale behind this choice is twofold: first, the top 30 cryptocurrencies by market cap generally exhibit higher liquidity, making them more suitable for market-making strategies. High liquidity is essential for minimizing latency and slippage, as low-liquidity environments can lead to significant delays in order execution and higher trading losses. Second, focusing on well-capitalized assets provides greater research value, as these assets are often more stable and widely traded, making the results of our study more relevant to real-world trading scenarios.

Coins without perpetual contracts were excluded from the universe. We visualize the correlation and clustering of these 30 cryptocurrencies to better understand their relationships, providing insight into how these assets co-move, which is crucial for both risk management and strategy calibration.

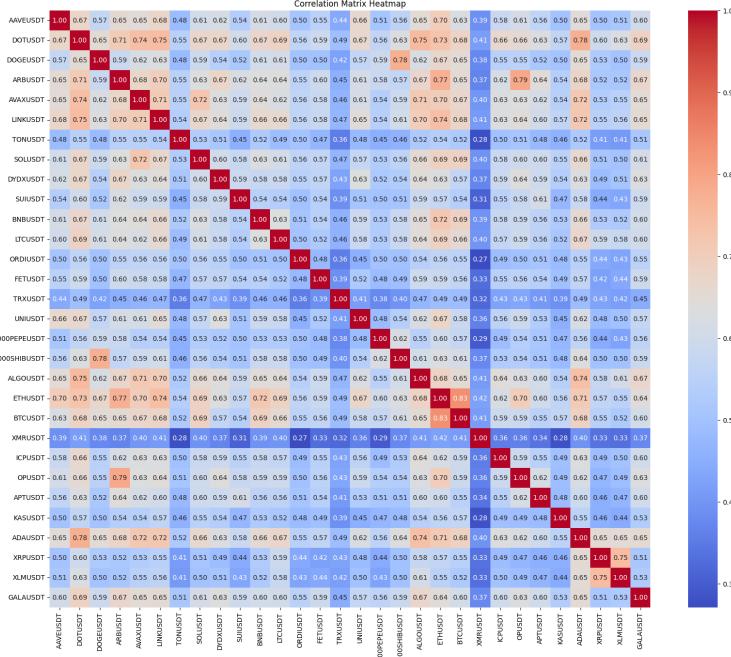


Figure 1: Correlation Matrix

### 2.1.1 Cryptocurrency Clustering Explanation

The clustering dendrogram and the accompanying classification table categorize the 30 selected cryptocurrencies into four distinct groups based on their market characteristics. This classification aids in the understanding of how different coins relate to each other and influences the design of our market-making strategy.

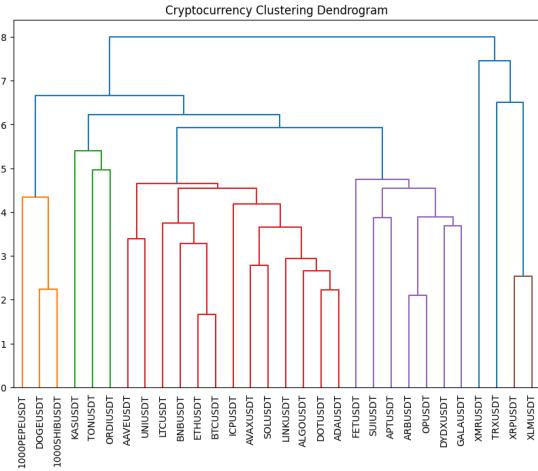


Figure 2: Clustering

#### 1. Layer-1 Protocols (Red):

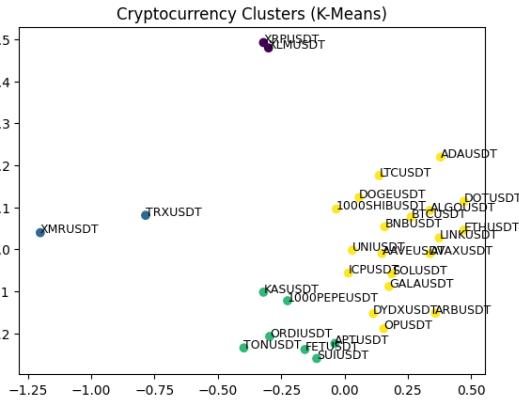


Figure 3: K-Means Cluster

- **Examples:** Ethereum (ETH), Bitcoin (BTC), Internet Computer (ICP), Solana (SOL)
- **Description:** Layer-1 protocols refer to the base layer blockchains that act as foundational platforms for decentralized applications (dApps) and other protocols. These coins often form the backbone of the cryptocurrency ecosystem due to their widespread adoption and usage. They are critical in facilitating transactions and smart contracts, making them highly liquid and relevant for long-term market-making strategies.

## 2. Meme Coins (Orange):

- **Examples:** Dogecoin (DOGE), Pepe (PEPE), Shiba Inu (SHIB)
- **Description:** Meme coins are a category of cryptocurrencies that often originate as internet jokes or cultural phenomena but have gained traction and liquidity over time. Despite their humorous beginnings, these coins have achieved significant market capitalization. Their volatility is typically higher compared to more established cryptocurrencies, which presents both opportunities and risks for market makers.

## 3. Decentralized Finance (DeFi) (Purple):

- **Examples:** Optimism (OP), dYdX (DYDX), Gala (GALA)
- **Description:** DeFi tokens are associated with decentralized financial applications, such as lending, borrowing, and trading platforms, without the need for traditional financial intermediaries. These coins are directly linked to the fast-growing DeFi sector, which is known for its innovative use cases but can also exhibit high levels of volatility due to rapid technological development and regulatory uncertainty.

## 4. Utility Tokens (Green):

- **Examples:** Kaspa (KAS), Toncoin (TON), Ordi (ORDI)
- **Description:** Utility tokens are digital assets that are used to access or power specific features of a blockchain platform or application. They often serve a functional purpose within their respective ecosystems, such as paying transaction fees or participating in governance. Utility tokens tend to have niche markets but can be less volatile than meme or DeFi tokens due to their underlying use cases.

## 5. Outliers (Gray):

- **Examples:** XRP, XLM
- **Description:** Certain cryptocurrencies, like XRP and XLM, don't fit neatly into the above categories. These outliers might represent unique protocols or cross-border payment solutions. Their behavior often diverges from the more common classifications, which may make it challenging to model within a standard market-making framework.

## 2.2 Candlestick Data

For our analysis, we utilized candlestick data rather than order book data. While order book data can offer a more granular view of market conditions, complete order book datasets often require costly subscriptions, which was beyond the scope of this paper. Additionally, the Hummingbot framework is designed to work seamlessly with candlestick data, making it a practical choice for developing and backtesting our market-making algorithms. Candlestick data provides key metrics such as open, high, low, and close prices for specific intervals, as well as the trading volumes, which are essential for identifying market trends and setting parameters for the algorithms. Figure 4 gives an example of our data.

BTCUSDT												
open_time	open	high	low	close	volume	close_time	quote_volume	count	taker_buy_volume	taker_buy_quote_volume	ignore	
1725148800000	58941.9	58941.9	58933.4	58933.8	41.719	1725148859999	2458844.5111	722	15.996	942775.9394	0	
1725148860000	58933.7	58940.4	58928.5	58928.5	17.332	1725148919999	1021460.2227	516	5.472	322482.8621	0	
1725148920000	58928.5	58945.1	58925.2	58945.1	15.938	1725148979999	939282.5365	542	10.222	602433.7369	0	
1725148980000	58945.0	58946.3	58926.7	58930.7	10.627	1725149039999	626320.4555	383	3.591	211653.7059	0	
1725149040000	58930.8	58951.4	58930.7	58951.4	10.291	1725149099999	606566.9869	442	9.081	535253.6374	0	

Figure 4: Data Example (First 5 Rows of BTC-USDT Candlestick Data)

## 2.3 Dataset Description

The dataset was collected from Binance and consists of multiple columns, each representing different aspects of market behavior. The main columns in the dataset include:

- **open time:** The timestamp marking the beginning of each one-minute interval.
- **open:** The price at the start of the interval.
- **high:** The highest price recorded during the interval.
- **low:** The lowest price recorded during the interval.
- **close:** The price at the end of the interval.
- **volume:** The total trading volume during the interval.
- **close time:** The timestamp marking the end of the interval.
- **quote volume:** The volume in terms of the quote asset (i.e., the asset that the base currency is traded against).
- **count:** The number of trades executed during the interval.
- **taker buy volume:** The volume of trades where the taker is the buyer.

- **taker buy quote volume:** The volume of taker-buy trades in the quote asset.
- **ignore:** A placeholder column that is not relevant to our analysis.

We collected one-minute candlestick data for 30 cryptocurrencies over a period of 45 days, specifically from September 1, 2024, to October 14, 2024. This resulted in approximately 60,000 data points for each cryptocurrency, forming the basis for our initial analysis and alpha generation.

To ensure the reliability of our analysis, we focused on cryptocurrencies with perpetual contracts and excluded those lacking sufficient historical data. For shorter time intervals such as 3-minute, 15-minute, and 1-hour candlestick data, we aimed to gather a similar volume of data, but in some cases, we were limited to around 30,000 data points due to the newer nature of certain coins.

## 3 Pure Market Making

Pure Market Making (PMM) is a trading strategy that involves the continuous placement of hedging order pairs, typically consisting of buy (long) limit orders and sell (short) limit orders. The sell orders are positioned above the buy orders, adhering to the fundamental principle of ‘buy low, sell high.’ At the core of this strategy are two key parameters: the reference price and the spread. The reference price represents the midpoint or average between the long and short order prices, serving as the anchor for order placement. The spread, defined as the distance between the buy and sell order prices, determines the level of price sensitivity and profitability of the strategy.

The effectiveness of PMM strategies depends heavily on the precise tuning of these parameters, which balance the trade-off between market risk and profitability. Baseline implementations of PMM strategies, such as PMM Simple and PMM Dynamic, have been provided by Hummingbot. These baseline strategies incorporate various approaches to parameter adjustment, offering practical frameworks for market making.

### 3.1 PMM simple

The PMM Simple strategy in Hummingbot has the fixed spread and static reference price, it allows users to configure a straightforward market-making bot, designed to maintain buy and sell orders on both sides of the order book.

### 3.1.1 Input parameters

- **Connector:** Specifies the exchange platform, such as KuCoin, OKX, and Binance, where the strategy will execute trades.
- **Trading Pair:** Indicates the market pair to trade, e.g., PEPE-USDT.
- **Leverage:** Defines the level of leverage to apply to the trades. Higher leverage amplifies both potential gains and risks.
- **Total Amount of Quote:** Sets the total value (in the quote currency, USDT in this case) to be utilized for the strategy.
- **Position Mode:** Chooses between HEDGE or ONE-WAY position mode. In hedge mode, the strategy can hold both long and short positions simultaneously.
- **Stop Loss Cooldown Time:** Determines the time interval (in minutes) the system will wait before reinitiating a stop loss process after an execution.
- **Executor Refresh Time:** Sets the time interval (in minutes) to refresh the executor's parameters and configurations.
- **Number of Buy Order Levels:** Defines how many buy orders to place at varying price levels.
- **Type of Spread Distribution:** Allows users to choose between manual or automatic distribution of the spread between order levels.
- **Spread for Levels:** Specifies the price spread between the current market price and the buy order for each level.
- **Amount for Levels:** Sets the amount to buy at each level.
- **Stop Loss:** Automatically closes a position when the asset's price drops by a certain percentage to limit losses.
- **Take Profit:** Automatically exits a position when the asset's price reaches a predetermined profit percentage.
- **Time Limit:** Sets a time-based exit from positions if neither the stop loss nor take profit conditions are met.
- **Trailing Stop:** Allows the system to automatically follow price movements upwards, closing the position if the price retraces by a set percentage (Delta).

### 3.1.2 Backtesting results

In our backtesting of the PMM Simple strategy, we focused on optimizing two critical parameters: Stop Loss Cooldown Time and Executor Refresh Time. We ran the backtest on the PEPE-USDT trading pair. The aim was to adjust these parameters in order to maximize the strategy's profit and loss (P&L).

1. **Stop Loss Cooldown Time:** This parameter defines the waiting period before the system reinitiates the stop-loss function after execution.
2. **Executor Refresh Time:** This parameter controls the frequency at which the strategy refreshes its configurations and checks for new opportunities in the market.

After conducting a series of backtests, we identified the optimal ranges for these two parameters that yielded the best P&L results during the backtesting period from the coordinate graph (Figure 5) and contour plot (Figure 6):

- **Stop Loss Cooldown Time:** The most effective range was between **8 to 9 minutes**. A cooldown period shorter than 8 minutes resulted in premature re-engagement of the stop loss mechanism, while a longer period reduced the system's responsiveness to volatile price movements.
- **Executor Refresh Time:** The best performance was achieved with an interval of **3 to 5 minutes**. Refreshing the executor too frequently (less than 3 minutes) caused unnecessary adjustments without significant price changes, while a slower refresh time (more than 5 minutes) led to missed opportunities for market re-entry after favorable price movements.

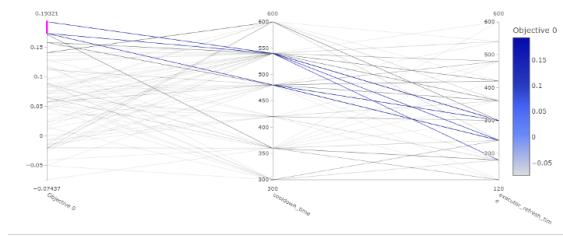


Figure 5: Parallel Coordinate Graph

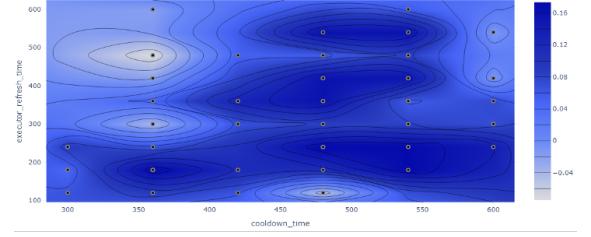


Figure 6: Contour Plot

## 3.2 PMM Dynamic

The PMM Dynamic Strategy utilizes a sophisticated approach to dynamically adjust both the mid-price and spread based on market conditions. The mid-price is adjusted using the

Moving Average Convergence Divergence (MACD) technical indicator, along with market volatility as measured by the Normalized Average True Range (NATR). The spread is also dynamically adjusted based on the NATR, reflecting the current market volatility.

### 3.2.1 Input parameters

The PMM Dynamic Strategy shares most of its parameters with the PMM Simple Strategy, but introduces several additional parameters to enhance its dynamic price and spread adjustments. These additional parameters include:

- **NATR Length:** Specifies the look-back period for calculating the **Normalized Average True Range (NATR)**, which measures market volatility. A higher NATR length captures longer-term volatility trends, while a lower value is more sensitive to recent price movements.
- **MACD Fast:** Defines the short-term exponential moving average (EMA) used in the **Moving Average Convergence Divergence (MACD)** indicator. This parameter influences how quickly the strategy reacts to price changes over shorter time periods.
- **MACD Slow:** Defines the long-term EMA for the MACD indicator. It smooths out price data over a longer period, providing a slower but more stable signal for mid-price adjustments.
- **MACD Signal:** Represents the signal line, which is the moving average of the MACD line. It is used to generate buy or sell signals when it crosses the MACD line, aiding the strategy in identifying market entry and exit points.

These additional parameters provide the strategy with the ability to adjust both mid-price and spreads based on momentum (MACD) and market volatility (NATR), allowing for more adaptive and efficient trading in volatile conditions.

### 3.2.2 Simulation Backtesting results

Same as PMM Simple, we tried to simulate the parameter combinations for the PEPE-USDT, and achieved a Sharpe ratio of about 0.86.

```

max_drawdown_pct -0.0588849517403619
max_drawdown_usd -58.98965011137752
net_pnl 0.11617606060746483
net_pnl_quote 116.17606060746482
profit_factor 1.1221817358617057
sharpe_ratio 0.8563748543464729

```

Figure 7: Backtesting Results



Figure 8: Hyperparameter Importance Analysis

## 4 Modeling

### 4.1 Alpha generation

#### 4.1.1 Pool of Alphas

To effectively translate candlestick data into actionable factors, we developed three distinct categories of indicators: candlestick characteristics, neighborhood trends, and volume-related factors to capture the key price movements. Each category aims to provide unique insights into market behavior, contributing to a comprehensive alpha-generation framework.

##### 1. Candlestick Characteristics

We constructed three indicators to capture the intrinsic features of candlesticks:

- (a) **Bar Portion with Direction:** This indicator ranges from -1 to 1, representing the proportion of the bar relative to the candlestick and its direction. For instance:
- An increasing candlestick is assigned a positive value.
  - A decreasing candlestick is assigned a negative value.
  - The magnitude reflects the significance of the move, with larger absolute values for more substantial changes.

$$\text{Bar Portion}_t = \frac{\text{Close}_t - \text{Open}_t}{\text{High}_t - \text{Low}_t}$$

- (b) **Bar Position:** This factor identifies the location of the bar within the candlestick, which is crucial for pattern recognition. For example:
- A candlestick resembling an inverse hammer (a small body near the top of the range with a long lower wick) is distinctly different from a hammer (the

inverse pattern). Without bar position, both patterns would appear identical in data, missing critical distinctions.

$$\text{Bar Position}_t = \frac{\frac{\text{Open}_t + \text{Close}_t}{2} - \text{Low}_t}{\text{High}_t - \text{Low}_t}$$

- (c) **Stick Length:** This indicator measures the high-low range of the candlestick, scaled by the Average True Range (ATR). By normalizing this measurement, we reduce the emphasis on insignificant small moves while highlighting more meaningful price changes.

$$\text{Stick Length}_t = \frac{\text{High}_t - \text{Low}_t}{\text{ATR}_{t-1}}$$

where ATR is Average True Range with a look-back period of 10.

## 2. Neighborhood Trends

To capture broader trends in price movements, we introduced two indicators:

- (a) **Slope:** The average rate of change in price over a defined neighborhood of data points. This captures the general upward or downward trajectory of the trend.

$$\text{Slope}_t = \frac{\text{Close}_t - \text{Close}_{t-1}}{\text{Close}_{t-1}}$$

- (b) **Curvature:** The rate of change in the slope, effectively measuring the acceleration or deceleration of the trend.

$$\text{Curvature}_t = \text{Slope}_t - \text{Slope}_{t-1}$$

These two indicators, when used together, provide a more nuanced understanding of market dynamics. For example, as demonstrated in the accompanying plot, slope identifies the direction of the trend, while curvature captures the intensity and turning points. While their effectiveness has been tested independently in this phase, our next phase will focus on combining them to uncover potential interactions.

## 3. Volume-Related Factors

To understand the role of trading volume in price movements, we developed three additional indicators:

- (a) **Change in Volume-Weighted Average Price (VWAP):** VWAP is a moving average weighted by volume, offering a price benchmark adjusted for trading activity. This indicator tracks changes in VWAP over time.

$$\text{VWAP} = \frac{\sum_{i=1}^n \text{TP}_i \cdot V_i}{\sum_{i=1}^n V_i}$$

$$\text{TP}_i = \frac{\text{High}_i + \text{Low}_i + \text{Close}_i}{3}$$

where V stands for Volume.

- (b) **VWAP Deviation from Close Price:** This measures the spread between VWAP and the actual close price, reflecting how far the closing price deviates from the volume-adjusted average. In the accompanying plot, this spread is represented by the difference between the blue (VWAP) and green (close price) lines.

$$\text{VWAP Deviation from Close Price}_t = \frac{\text{VWAP}_t - \text{Close}_t}{\text{Close}_t}$$

- (c) **Buy Volume Portion:** This indicator estimates the relative strength of buyers versus sellers by measuring the proportion of buy volume compared to total trading volume.

$$\text{Buy Volume Portion}_t = \frac{\text{Taker Buy Volume}_t}{\text{Volume}_t}$$

While each indicator was tested separately, our future work will focus on integrating these factors to explore their interactions. We anticipate that combining these indicators will uncover deeper insights into market behavior and improve the predictive power of our alpha-generation models.

#### 4.1.2 Quintile Analysis

After generating our signals, we compared their performance to the MACD signal used in the original PMM Dynamic strategy to evaluate their effectiveness. Our primary evaluation method was quintile analysis, a technique to assess whether a given signal contains alpha by analyzing its relationship with asset performance. Below, we provide a detailed breakdown of this process and the insights derived from our findings.

##### 1. Introduction to Quintile Analysis

Quintile analysis divides a signal into five equal groups, or ‘quintiles,’ to evaluate the relationship between the signal and asset performance. For each quintile, which ranges from the lowest 20% to the highest 20% of signal values, we calculate the average returns. Visualizing these averages helps identify patterns and relationships, such as monotonic trends, which may indicate the predictive power (or alpha) of a signal. Specifically:

- A **monotonically increasing pattern** suggests that higher signal values correspond to higher returns.
- A **monotonically decreasing pattern** indicates that lower signal values correspond to higher returns.

This method provides a clear way to assess whether a signal offers consistent and predictable insights into asset behavior.

## 2. MACD Signal Analysis

We began by performing the quintile analysis on the MACD signal implemented in PMM Dynamic, testing its relationship with returns across all coins in our trading universe. As shown in Figure 9, there was no clear linear relationship between the MACD signal and returns. This lack of consistency indicates that the MACD signal does not exhibit strong alpha potential. As a result, we chose not to further explore or modify this signal and shifted our focus to analyzing the signals we generated.

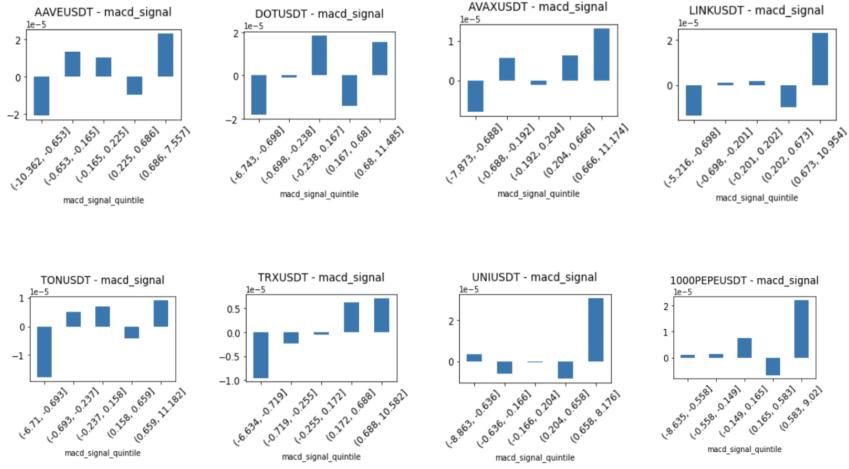


Figure 9: MACD Signal Quintile Analysis

## 3. Analysis of Generated Signals

We conducted the quintile analysis on 10 custom-generated signals, aiming to identify those with strong, consistent patterns. Among these, five signals demonstrated clear monotonic relationships—either increasing or decreasing—that were largely consistent across our trading universe. For example:

- The **Bar Portion signal** (bar\_portion) exhibited a monotonically decreasing pattern.
- The **VWAP-to-Close Price Difference signal** (diff\_vwap\_close) showed a monotonically increasing trend.

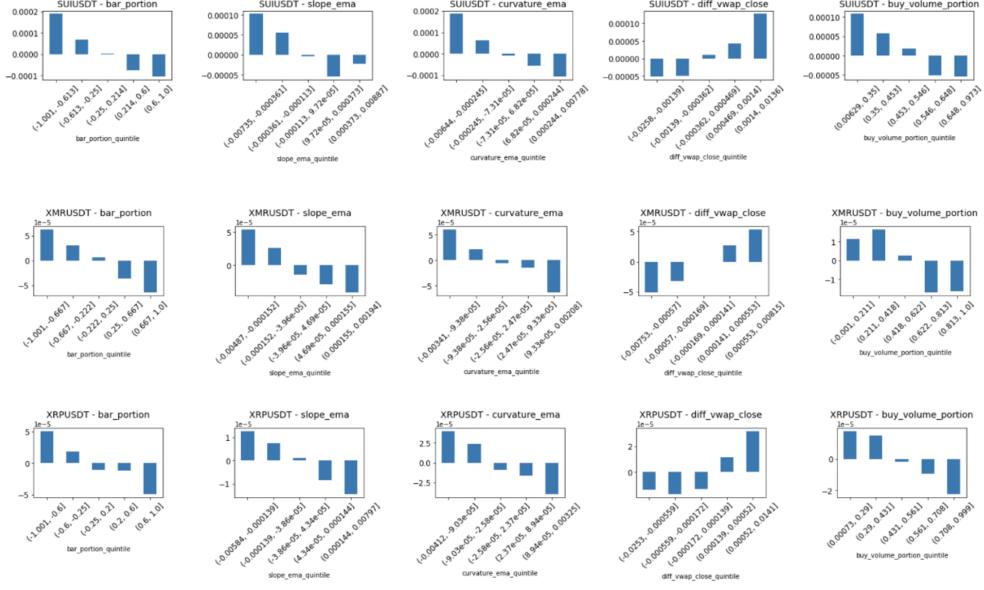


Figure 10: Crafted Alphas Quintile Analysis

#### 4. Quantifying Monotonicity

To avoid relying solely on visual impressions, we quantified the monotonicity of each signal by calculating the proportion of consistent monotonic behavior across all coins in our trading universe. Based on this analysis, the top-performing signals were:

- **Bar Portion:** Monotonic proportion of 73%
- **Curvature EMA:** Monotonic proportion of 73%

as shown in Table 11.

	increasing	decreasing
<b>bar_position</b>	<b>0.10</b>	<b>0.17</b>
<b>bar_portion</b>	<b>0.00</b>	<b>0.73</b>
<b>stick_length</b>	<b>0.27</b>	<b>0.03</b>
<b>slope_sma</b>	<b>0.00</b>	<b>0.43</b>
<b>slope_ema</b>	<b>0.00</b>	<b>0.40</b>
<b>curvature_sma</b>	<b>0.00</b>	<b>0.57</b>
<b>curvature_ema</b>	<b>0.00</b>	<b>0.73</b>
<b>vwap_change</b>	<b>0.00</b>	<b>0.13</b>
<b>diff_vwap_close</b>	<b>0.40</b>	<b>0.00</b>
<b>buy_volume_portion</b>	<b>0.00</b>	<b>0.53</b>

Figure 11: Monotonicity Analysis

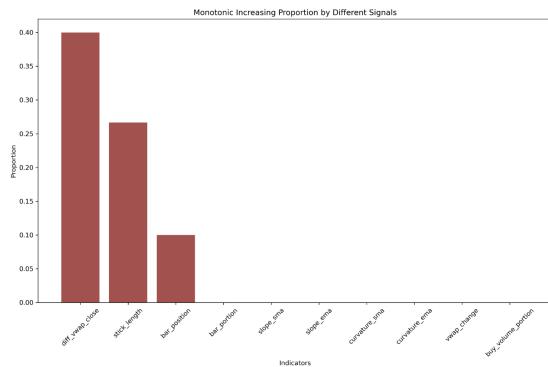


Figure 12: Monotonic Increasing Proportion by Different Signals

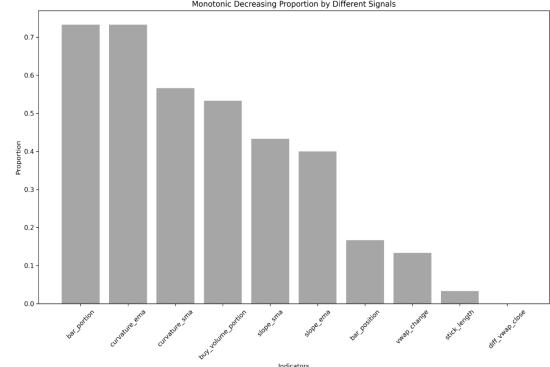


Figure 13: Monotonic Decreasing Proportion by Different Signals

For the remainder of this section, we focused primarily on the Bar Portion signal, as it exhibited both strong monotonicity and clear interpretability.

## 5. Bar Portion Signal: Intuition and Insights

The Bar Portion signal measures the proportion of the bar within a candlestick and is calculated as:

This value ranges from -1 to 1:

- **-1:** Indicates a price decrease within the tick (Open = High, Close = Low).
- **1:** Indicates a price increase within the tick (Open = Low, Close = High).
- Values near **0:** Represent minimal price movement within the tick.

## 6. Key Insights from Quintile Analysis

From the quintile analysis graph for the Bar Portion signal, we observed the following patterns:

- Decreasing Returns with Increasing Bar Portion:** Average returns decrease as Bar Portion values rise. This suggests that significant price increases within a tick are often followed by large declines in the next tick, and vice versa.
- Mean Reversion Behavior:** Large movements within a tick frequently lead to reversals in the following tick.
- Stability for Middle Quintiles:** For mid-range Bar Portion values (40% - 60%), minimal price changes within a tick often correspond to stable prices in the subsequent tick.

These behavioral patterns highlight the potential of the Bar Portion signal to capture market tendencies, particularly in identifying mean-reverting opportunities. The consistency of these insights across multiple coins underscores the robustness of this signal as a predictive tool for market-making strategies.

### 4.1.3 Single Alpha Backtesting

The primary goal at this stage was to enhance the PMM Dynamic Strategy by improving the reference price mechanism. Specifically, we aimed to identify a superior indicator from our pool of alphas to replace the original MACD signal. To evaluate the effectiveness of these potential replacements, we conducted baseline backtesting using simple directional trading strategies.

#### 1. Baseline Backtesting Methodology

Building on the results of the quintile analysis, which validated the monotonicity and linear relationship of certain alphas, we implemented a baseline backtesting framework. This methodology employed rolling one-dimensional linear regression to quantify the average impact of each potential alpha on subsequent price changes. The backtesting process used the following approach:

- (a) **Rolling Window Setup:** The regression model was trained using a 36-day window, equivalent to 51,840 records of minute-level data. Predictions for the next 9 days (12,959 records) were generated iteratively based on the calculated coefficients.
- (b) **Directional Strategy:** The strategy maintained a 100% position in either a long or short direction. A positive prediction signaled a long position, while a negative prediction resulted in a short position.
- (c) **Portfolio Construction:** To ensure a robust evaluation, the balance was equally distributed across all 30 coins in the trading universe. This setup effectively simulated 30 independent portfolios running the same strategy in parallel. The aggregate balance, calculated as the sum of all portfolio balances, was used to evaluate performance:

$$\text{Balance}_t = \sum_{i=1}^{30} P_{i,t}, \quad \text{where } P_{i,t} \text{ represents the balance for coin } i \text{ at time } t.$$

$$\text{Return}_t = \frac{\text{Balance}_t - \text{Balance}_{t-1}}{\text{Balance}_{t-1}}, \quad \text{with } \text{Balance}_0 = 10,000.$$

This methodology mitigated risk and bias by diversifying across multiple coins and time intervals, reducing the influence of idiosyncratic factors.

#### 2. Comparison with MACD Signal

Among the alphas tested, Bar Portion emerged as the best-performing signal, consistent with the quintile analysis results. Figures 14 and 15 compare the cumulative returns of MACD and Bar Portion over a 9-day period.



Figure 14: MACD Baseline P&L

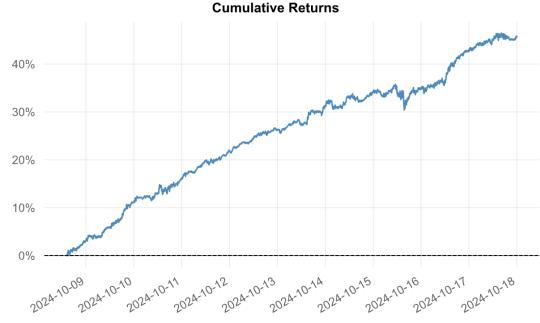


Figure 15: Bar Portion Baseline P&L

#### Performance Metrics:

- (a) **Return:** Bar Portion achieved an accumulative return of 45.84%, while MACD delivered a negative return of -0.59%.
- (b) **Maximum Drawdown:** Bar Portion demonstrated lower drawdowns at 3.94%, compared to 8.71% for MACD.
- (c) **Sharpe Ratio:** Bar Portion achieved a Sharpe ratio of 0.78, significantly outperforming MACD's -0.01.

	<b>MACD</b>	<b>Bar Portion</b>
<b>Sharpe</b>	<b>-0.01</b>	<b>0.78</b>
<b>Accumulative Return (%)</b>	<b>-0.59</b>	<b>45.84</b>
<b>Sortino</b>	<b>-0.01</b>	<b>1.14</b>
<b>Maximum Drawdown (%)</b>	<b>-8.71</b>	<b>-3.94</b>
<b>Annualized Volatility (%)</b>	<b>0.94</b>	<b>0.9</b>

Figure 16: Baseline MACD vs. Bar Portion Performance Metrics Table

The results indicate that Bar Portion provides more consistent and stable returns compared to MACD, as illustrated in Table 16. This highlights its potential as a replacement reference price indicator in the PMM strategy.

### 3. Discussion

While Bar Portion demonstrated strong performance during baseline backtesting, certain limitations and considerations need to be addressed:

- (a) **Trading Costs:** The backtesting framework did not account for transaction costs. High-frequency rebalancing at one-minute intervals could lead to significant discrepancies between backtested and real-world performance due to elevated trading fees.

- (b) **Directional Focus:** The methodology focused solely on directional trading, whereas our ultimate objective is market-making or hedging. This directional emphasis might shift the strategy's focus toward profiting from large price movements, which diverges from the ideal of earning from non-directional market fluctuations.

Despite these limitations, the backtesting process effectively validated the predictive power of Bar Portion as a reference price indicator, laying the groundwork for its integration into a refined market-making strategy.

## 4.2 Risk Management

Risk management plays a critical role in the design of any market-making algorithm, especially in the volatile cryptocurrency market. We focus on implementing the triple barrier strategy to manage risk across different crypto assets. The triple barrier strategy is designed to close positions based on three key parameters: stop loss, take profit, and the time barrier, effectively controlling risk and preserving capital during adverse market movements.

### 4.2.1 Triple Barrier Strategy

Several parameters influence the effectiveness of the triple barrier strategy:

- **Stop Loss:** This defines the maximum loss that can be tolerated on a trade before the position is automatically closed. It helps limit downside risk, especially in highly volatile environments like cryptocurrency markets.
- **Take Profit:** The take profit level specifies the price at which a trade is exited once it has reached a predetermined profit target. This ensures that profits are secured when favorable market conditions arise.
- **Spread:** The difference between the bid and ask prices plays a significant role in determining when and where orders are placed. A wider spread can result in fewer trades but lower the chance of immediate losses, while a narrower spread can increase trade frequency but expose the strategy to more market fluctuations.
- **Refresh Time:** This parameter controls how often the algorithm recalculates and updates the orders based on current market conditions. More frequent refresh times can help capture short-term price movements, but they also increase trading costs and latency issues.
- **Trailing Stop:** Unlike a static stop loss, a trailing stop moves with the market when the price moves in a favorable direction, locking in profits while still allowing for further gains. It's an adaptive risk control measure that adjusts as market prices evolve.

#### 4.2.2 Calibration of Parameters Across Different Coins

Calibrating these parameters for different coins is crucial, as each cryptocurrency has its own unique volatility, liquidity, and trading patterns. A one-size-fits-all approach may not work across the diverse set of assets in the crypto universe:

- **Volatility-Based Calibration:** Highly volatile coins may require wider stop-loss and take-profit levels to prevent premature exits due to normal price fluctuations. Lower-volatility coins, on the other hand, can function with tighter parameters, as their price movements are more stable and predictable.
- **Liquidity Consideration:** For highly liquid coins, we can afford to use narrower spreads and shorter refresh times, as there is less risk of slippage and order execution delays. For low-liquidity coins, a wider spread may be necessary to protect against sudden price swings and ensure that the algorithm avoids entering into trades that cannot be efficiently executed.
- **Historical Performance:** Backtesting the strategy with different parameter settings on historical data is essential to understanding how each coin responds to various market conditions. The stop loss, take profit, and trailing stop levels can be adjusted based on historical volatility and price patterns specific to each asset.
- **Risk Tolerance:** The calibration should also consider the risk tolerance for each coin. For instance, riskier assets might warrant tighter risk controls, while more stable coins could tolerate a more aggressive trading approach with looser risk parameters.

#### 4.2.3 Parameters Optimization Results

In this study, we focus on optimizing key parameters of a pure market-making model to manage risk effectively in cryptocurrency trading. Leveraging Optuna, a powerful hyper-parameter optimization framework, we analyzed the interaction between crypto monthly volatility and four critical parameters: *start spread*, *step spread*, *stop loss*, and *take profit*. These parameters play pivotal roles in determining the strategy's performance.

- **Start spread:** Defines the initial order placement distance from the mid-price.
- **Step spread:** Specifies the incremental distances between successive order levels.
- **Stop loss:** Establishes thresholds to limit potential losses.
- **Take profit:** Sets thresholds to secure realized profits.

Through backtesting cryptocurrency data for October 2024, we employed a two-level spread structure in the market-making model to assess how varying volatility environments influence

the performance of these parameters. Our data points represent the optimized values for *start spread*, *step spread*, *stop loss*, and *take profit* for each cryptocurrency in our universe of 30 assets, obtained by maximizing the Sharpe ratio. This approach ensures that the selected parameter values align with the objective of achieving an optimal balance between return and risk, as captured by the Sharpe ratio.

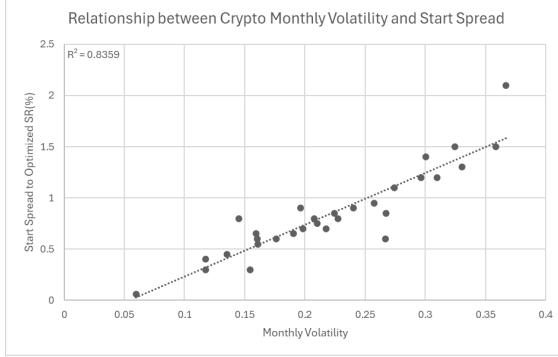


Figure 17: Relationship between Crypto Monthly Volatility and Start Spread

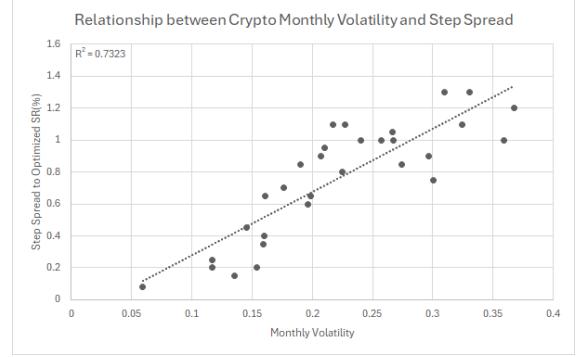


Figure 18: Relationship between Crypto Monthly Volatility and Step Spread

The relationships between monthly cryptocurrency volatility and the *start spread* and *step spread* were examined using linear regression. As illustrated in Figures 17 and 18, our empirical results indicate a strong positive correlation between volatility and these spread parameters. The optimized values for both *start spread* and *step spread* are approximately 4 to 5 times the observed monthly volatility. This relationship underscores an intuitive principle: more volatile markets necessitate wider spreads to balance risk and profitability.

The robustness of these findings is evidenced by the high  $R^2$  values, which highlight the linear trends' statistical significance. These results provide a practical benchmark for calibrating spreads in live trading environments, enabling market makers to adapt their strategies dynamically to volatility shifts.

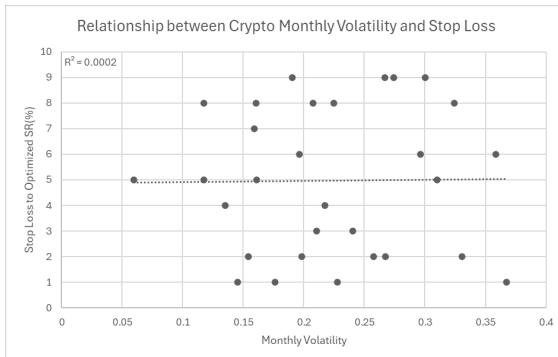


Figure 19: Relationship between Crypto Monthly Volatility and Stop Loss

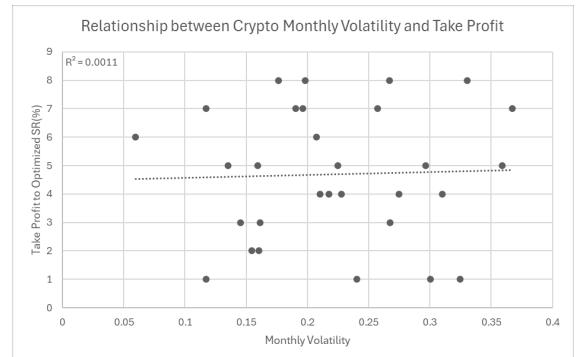


Figure 20: Relationship between Crypto Monthly Volatility and Take Profit

In contrast, the analysis of *stop loss* and *take profit* thresholds yielded markedly different results. Figures 19 and 20 reveal no discernible correlation between monthly cryptocurrency volatility and the optimized values of these parameters. The near-zero  $R^2$  values suggest that *stop loss* and *take profit* thresholds are not significantly influenced by market volatility.

This lack of correlation implies that these parameters are better optimized independently of volatility metrics. By leveraging Optuna’s optimization capabilities, we calibrated the thresholds based on specific market conditions to achieve more favorable trading outcomes. While the absence of a clear volatility-driven relationship limits predictive adjustments for these parameters, the flexibility of Optuna’s framework ensures that they remain adaptable to diverse trading scenarios.

Our findings highlight the importance of tailoring parameter optimization to the unique characteristics of each input. While *start spread* and *step spread* exhibit strong volatility-dependent relationships, *stop loss* and *take profit* require independent optimization for effective risk management. This distinction underscores the value of a hybrid approach that integrates both volatility-driven benchmarks and data-driven optimization tools in the development of robust market-making strategies.

## 4.3 Live Trading

### 4.3.1 Objective

The goal of live trading was to evaluate the performance of the PMM Dynamic and PMM BP strategies across three different cryptocurrency clusters: Layer 1 Protocols, Meme Coins, and Decentralized Finance (DeFi). Specifically, we selected one coin from each cluster—GALA, DOGE, and SOL—for a 24-hour trading period.

### 4.3.2 Setup

1. Three OKX accounts were used to conduct the trading.
2. Each account was dedicated to running a single cryptocurrency pair using one of the two strategies: PMM Dynamic or PMM BP.
3. The selected pairs for the analysis were:
  - GALA-USDT
  - DOGE-USDT
  - SOL-USDT

These pairs were chosen to represent different range of volatility and categories.

4. All three pairs were traded concurrently for approximately 24 hours to ensure uniform market conditions during the test period.
5. Due to technical constraints, each pair could only be run on one device at a time. To address this single-device limitation, two team members were assigned to manage the same pair on separate devices.
6. As a result, the trading periods for each pair were not perfectly synchronized but overlapped sufficiently to ensure meaningful comparisons.

#### 4.3.3 Performance Metric

- **Net P&L:** Represents the total profit or loss generated during the trading period, providing a measure of each strategy's performance both in absolute terms and as a percentage.

#### 4.3.4 Challenges and Limitations

- **Device Constraints:** The incapability to run multiple pairs on a single device led to slight variations in trading times.
- **Synchronization Issues:** Since trading sessions for the same pair were managed on different devices and by different team members, minor discrepancies in trading durations were unavoidable.

## 5 Results & Discussion

### 5.1 Backtesting vs. Live Trading

We compared the performance of backtesting versus live trading for the PMM BP strategy applied, for example, to the SOL-USDT trading pair over a 24-hour time horizon.

#### 5.1.1 Setup

Both backtesting and live trading were conducted with a total amount quote of 500 USDT, and the same risk management settings were applied to ensure consistency across both testing environments.

### 5.1.2 Results

- **Backtesting:** The strategy resulted in a net loss of -\$0.69 USDT, equating to -0.14% in PnL, with a total trading volume of approximately \$34,000.
- **Live Trading:** The strategy showed a positive net PnL of \$1.31, or 0.26%, despite a lower total trading volume of \$24,000.

	Net P&L (\$)	Net P&L (%)	Total Volume (\$)
Backtesting	-0.69	-0.14	33982.37
Live Trading	1.31	0.26	23922.50

Figure 21: Backtesting vs. Live Trading Performance Comparison Table



Figure 22: SOL-USDT Backtesting



Figure 23: SOL-USDT Live Trading

### 5.1.3 Analysis of Discrepancy

The differences in performance between backtesting and live trading can be attributed to several factors:

1. **Time Horizon Alignment:** The start and end times did not perfectly align at exact 12:00 AMs, leading to slight variations in the results.
2. **Trading Costs:** During backtesting, a trading cost of 0.06% was implemented, while in live trading, the actual cost was  $2 \times 0.02\% = 0.04\%$  without accounting for slippage. The conservative trading cost assumption in backtesting was deliberate and aligns with best practices for strategy evaluation.

These results emphasize that while backtesting provides a controlled and consistent environment, live trading captures real market dynamics, including volatility and slippage, which can significantly impact performance. Despite the lower trading volume, the live trading results outperformed backtesting, demonstrating the importance of real-world testing for accurate strategy assessment.

## 5.2 Live Trading: PMM Dynamic vs. PMM BP

This section presents a detailed comparison of the live trading performance of the PMM BP and PMM Dynamic strategies, focusing on key metrics such as profitability, liquidity, and efficiency. By applying these strategies to the SOL-USDT, DOGE-USDT, and GALA-USDT trading pairs over a 24-hour period, the objective was to identify their respective strengths and weaknesses under real market conditions. The results are analyzed based on trading volumes, realized profits, and risk management indicators, providing a comprehensive understanding of how these strategies perform in practice.

### 5.2.1 Setup

As discussed earlier, the live trading evaluation focuses on three trading pairs: SOL-USDT, DOGE-USDT, and GALA-USDT. To ensure a fair comparison between the PMM BP and PMM Dynamic strategies, all key trading parameters, including total quote amount, leverage, stop-loss, take-profit thresholds, and other relevant settings, were kept identical across both strategies. This standardization isolates performance differences to the strategies themselves.

### 5.2.2 Results

To compare the performance of PMM Dynamic and PMM BP, results are analyzed across different cryptocurrencies. The analysis focuses on key metrics, including profitability, liquidity provision, and operational efficiency, to evaluate the general effectiveness of each strategy. Additionally, a vertical comparison is conducted to assess performance under varying market conditions, providing a comprehensive evaluation of the strategies' robustness and adaptability.

1. **SOL-USDT:** Based on Figure 31 and Figure 32, we make the following conclusions for SOL-USDT in the live trading setting.
  - **Profitability:** The **PMM BP strategy** demonstrated a modest profit **\$1.214** compared to the **PMM Dynamic strategy**, which incurred a loss **\$-1.593**. Therefore, PMM BP outperformed in terms of positive returns.

- **Liquidity and Risk Management:** PMM Dynamic had higher liquidity placed **\$6,250.58** compared to PMM BP **\$1,861.80**, indicating a more capital-intensive approach. These two strategies minimized unrealized PnL, demonstrating the effective risk management process.
- **Efficiency:** Overall, **PMM BP** achieved better profitability with higher trading volume **\$23,922.5** and lower liquidity requirements. However, for **PMM Dynamic**, while it placed significantly higher liquidity, this strategy failed to convert this into net gains.
- **Conclusion:** This shows that the PMM BP strategy was more effective than the PMM Dynamic. Additionally, the PMM BP showed higher efficiency in percentage returns.

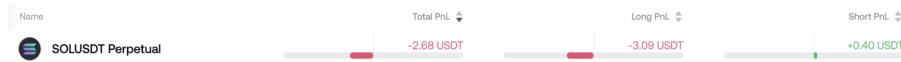


Figure 24: Example SOL’s PnL from PMM Dynamic

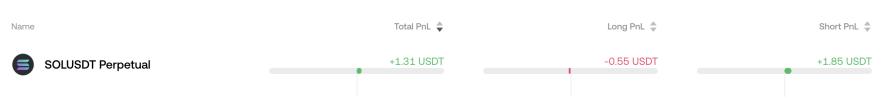


Figure 25: Example SOL’s PnL from PMM BP

2. **DOGE-USDT:** From Figure 33 and Figure 34, the following observations are summarized as below.

- **Profitability:** Both strategies achieved similar absolute profits, **\$3.191** for PMM BP and **\$3.216** for PMM Dynamic, with PMM Dynamic slightly outperforming. Besides, PMM BP achieved a higher percentage return **0.249%** relative to trading volume, indicating a more efficient capital utilization.
- **Liquidity and Risk Management:** For DOGE, PMM Dynamic placed significantly higher liquidity **\$-1,133.85** compared to PMM BP **\$-81.98**, reflecting a more capital-intensive approach. Both strategies effectively minimized unrealized PnL, showcasing a robust risk management process in live trading conditions.
- **Efficiency:** PMM Dynamic was more aggressive, achieving a higher trading volume **\$4,733.85** compared to PMM BP **\$1,281.98**. This might be the reason that the PMM Dynamic achieved a higher profit.
- **Conclusion:** In general, the results show that the PMM BP strategy has a comparable performance with PMM Dynamic.

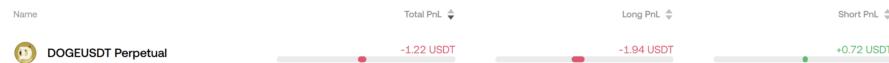


Figure 26: Example DOGE’s PnL from PMM Dynamic



Figure 27: Example DOGE's PnL from PMM BP

3. **GALA-USDT**: Figure 29 and Figure 28 come to the following conclusions.

- **Performance:** From the PnL plot, PMM BP consistently incurred lower total losses compared to PMM Dynamic, making it the preferable strategy in this context.
- **A Closer Look:** Both strategies experienced significant losses when trading the cryptocurrency Gala during the specified time interval. To analyze the underlying causes, a deeper investigation was conducted. Figure 30 provides an overview of Gala's market dynamics during the live trading period. The analysis revealed that Gala's price exhibited multiple sharp increases followed by gradual declines. Consequently, short limit orders were opened and remained in unfavorable positions as the price continued to rise, leading to suboptimal outcomes. This explains the observed disparity, where long positions outperformed short positions, aligning intuitively with the market's behavior.
- **Conclusion:** In conclusion, while further enhancements in volatility adjustments under varying market conditions are necessary, the new BP alpha strategy demonstrated improved profitability, indicating its effectiveness as a positive modification.



Figure 28: Example GALA's PnL from PMM Dynamic



Figure 29: Example GALA's PnL from PMM BP



Figure 30: Gala Live Trading Candlestick Chart

## 6 Conclusion

We successfully developed and evaluated automated market-making algorithms for cryptocurrency perpetual contracts using the Hummingbot platform. Leveraging data-driven techniques and robust risk management strategies, the PMM Dynamic model was designed and refined to address the unique challenges of the crypto market.

Extensive data exploration, including candlestick data research and correlation analysis in a diverse cryptocurrency universe, ensured that the strategies were grounded in empirical insights. Alpha-generation techniques and systematic backtesting provided a solid foundation for optimizing key parameters such as spread and risk thresholds. Indicators like Bar Portion were identified as effective tools for capturing market tendencies, enhancing predictive capabilities.

The integration of the triple barrier strategy highlighted the importance of risk management in volatile markets, striking a balance between profit realization and loss mitigation. Parameter optimization using tools like Optuna emphasized the necessity of tailoring strategies to specific market conditions.

Live trading results validated the practical viability of the proposed models, demonstrating notable improvements over backtesting in terms of profitability and risk control. The comparison between PMM Dynamic and PMM BP strategies offered insight into the nuances of strategy performance under real-world conditions.

This work contributes to the growing field of algorithmic trading in cryptocurrency markets by presenting adaptable, data-driven market-making frameworks. Future research could build upon these strategies by selecting appropriate coin spaces, analyzing cross-frequency

candlesticks, and exploring multi-strategy integrations to further optimize performance.

## 7 Appendix

### 7.1 Dashboard Figures

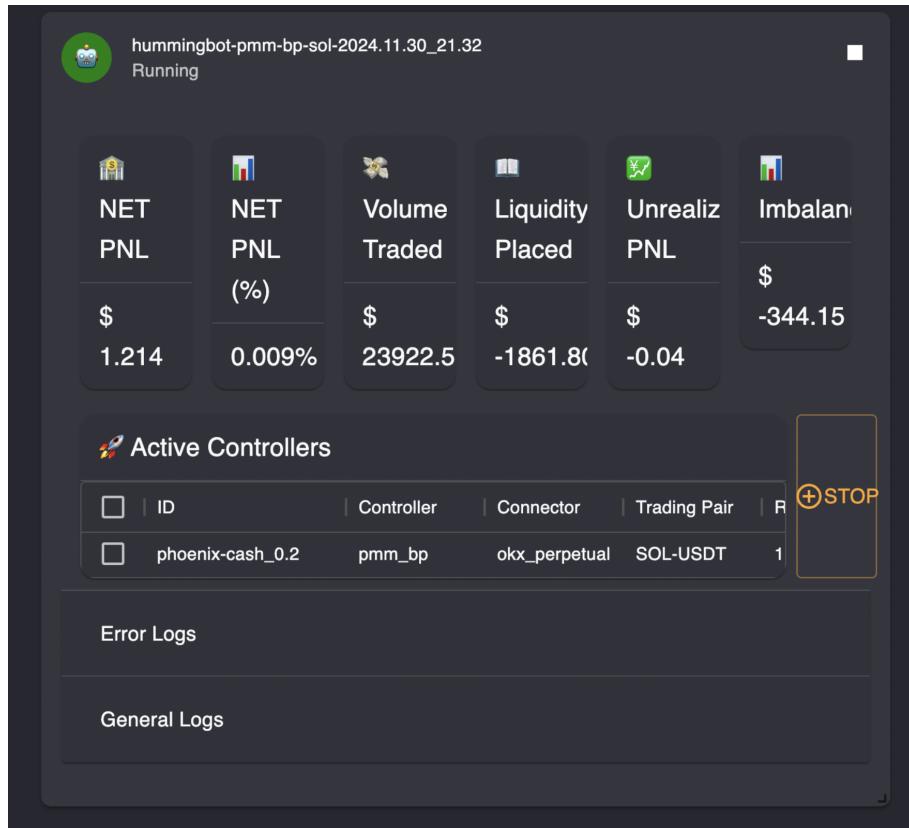


Figure 31: PMM BP for SOL in Live Trading

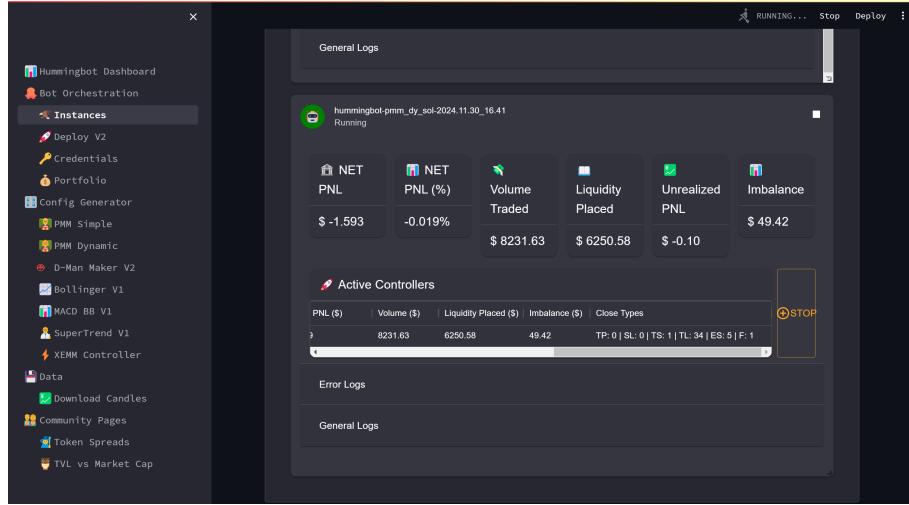


Figure 32: PMM Dynamic for SOL in Live Trading

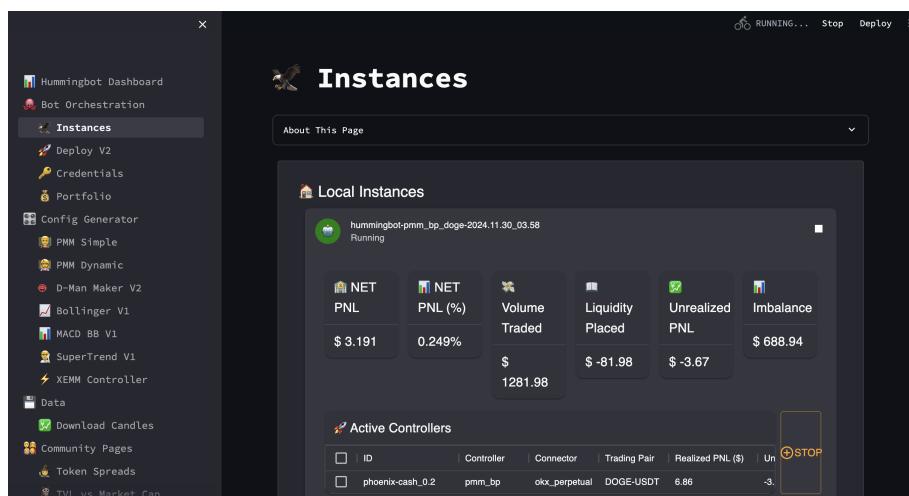


Figure 33: PMM BP for DOGE in Live Trading

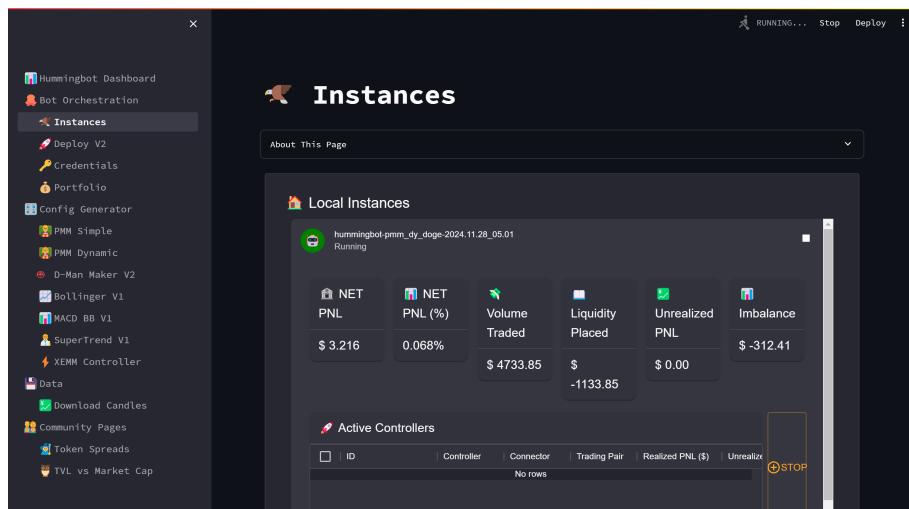


Figure 34: PMM Dynamic for DOGE in Live Trading

## References

- [AS08] Marco Avellaneda and Sasha Stoikov. High frequency trading in a limit order book. *Quantitative Finance*, 8:217–224, 04 2008.
- [CBW21] N. Crone, E. Brophy, and T. Ward. Exploration of algorithmic trading strategies for the bitcoin market. *arXiv preprint*, arXiv:2110.14936, 2021.
- [Dai18] Yiwen Dai. The rise and latest research progress of algorithmic trading. *Shenzhen Stock Exchange Research*, 2018.
- [HZY23] J. He, C. Zheng, and C. Yang. Integrating tick-level data and periodical signal for high-frequency market making. *arXiv preprint*, arXiv:2306.17179, 2023.
- [KB18] P. Katsiampa and S. Bekiros. Take profit and stop loss trading strategies comparison in algorithmic trading. *Journal of Risk and Financial Management*, 11(3):56, 2018.
- [KL23] J. Kim and S. Lee. P-triple barrier labeling: Unifying pair trading strategies and triple barrier method. *Journal of Financial Engineering*, 10(2):123–145, 2023.
- [Lat21] A. Latif. Developing a fully automated trading algorithm for the cryptocurrency trading market. Master’s thesis, LUT University, 2021.
- [Mon24] M. Monga. Decentralised finance and automated market making. *SSRN*, 2024.
- [SHJB22] Isabela Ruiz Roque da Silva, Eli Hadad Junior, and Pedro Paulo Balbi. A review of cryptocurrency trading algorithms. *Journal of Forecasting*, 41(5):1024–1045, 2022.