RESEARCH-ARTICLE

Attention-Based Multi-Asset Order Flow Networks for Enhanced Mid-Price Prediction

HAMIDREZA BANDEALINAEINI, Sharif University of Technology, Tehran, Tehran, Iran

MOHAMMAD SHARIFKHANI, Sharif University of Technology, Tehran, Tehran, Iran

ERFAN SALAVATI, Amirkabir University of Technology, Tehran, Tehran, Iran

Published: 15 November 2025

Citation in BibTeX format

ICAIF '25: 6th ACM International Conference on AI in Finance
November 15 - 18, 2025
Singapore, Singapore

ICAIF '25: Proceedings of the 6th ACM International Conference on AI in Finance (November 2025)
https://doi.org/10.1145/3768292.3770430
ISBN: 9798400722202

Attention-Based Multi-Asset Order Flow Networks for Enhanced Mid-Price Prediction

Hamidreza Bandealinaeini
Sharif University of Technology
Tehran, Tehran, Iran
hamidreza.bandealinaeini@ee.sharif.edu

Mohammad Sharifkhani
Sharif University of Technology
Tehran, Tehran, Iran
msharifk@sharif.edu

Erfan Salavati
Amirkabir University of Technology
Tehran, Tehran, Iran
erfan.salavati@aut.ac.ir

Abstract

Financial markets are complex adaptive systems where assets continuously influence each other through dynamic and nonlinear interactions. Accurate mid-price forecasting in high-frequency trading (HFT) depends on capturing these market-wide dynamics. While order flow imbalance (OFI) features have proven more effective than raw limit order book (LOB) data for short-term forecasting, most existing models remain limited to single-asset dynamics, ignoring informative signals from related instruments.

We propose OF-MATNet, a deep learning-based multi-asset forecasting framework that leverages OFI data from multiple Nasdaq-listed assets. Our approach captures nonlinear cross-asset dependencies through a Transformer-based architecture with multi-axis attention mechanisms over time, assets, and order book levels, enhanced with positional encoding. Informative peer assets for each target are selected using rolling-window Granger causality tests conducted during the training phase, enabling the model to exploit statistically validated cross-asset influences.

Experiments on 110 assets show that OF-MATNet significantly outperforms both our single-asset baseline (OF-SATNet) and state-of-the-art models such as DeepLOB, BINCTABL, and TLOB. OF-MATNet achieves consistent R^2 improvements in over 90% of cases, with larger gains for assets highly influenced by peers or at longer prediction horizons. Further analysis reveals that temporal attention contributes most to forecasting, but cross-asset and level-wise information are critical in enhancing accuracy. These findings underscore the practical value of modeling nonlinear cross-asset relationships for strategic financial decision-making.

CCS Concepts

• **Computing methodologies** → **Neural networks**; • **Applied computing** → **Forecasting**; • **Networks** → *Network design principles.*

Keywords

Order Book, Order Flow Imbalance, Mid-Price, High Frequency Trading

ICAIF '25, Singapore, Singapore

ACM ISBN 979-8-4007-2220-2/25/11
https://doi.org/10.1145/3768292.3770430

ACM Reference Format:
Hamidreza Bandealinaeini, Mohammad Sharifkhani, and Erfan Salavati. 2025. Attention-Based Multi-Asset Order Flow Networks for Enhanced Mid-Price Prediction. In *6th ACM International Conference on AI in Finance (ICAIF '25), November 15–18, 2025, Singapore, Singapore.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3768292.3770430

1 Introduction

Financial markets operate as complex adaptive systems, where the actions of traders and institutions across different assets create interconnected dynamics. In high-frequency trading (HFT), accurately forecasting short-term mid-price movements—the average of the best bid and ask prices—is crucial for applications such as market making, liquidity provision, and execution optimization. Given the granularity and speed of modern electronic markets, even marginal improvements in predictive accuracy can result in meaningful economic gains.

A key breakthrough in modeling price formation was introduced by Cont, Kukanov, and Stoikov [9], who proposed the concept of order flow imbalance (OFI) as a primary driver of short-term price changes. OFI captures the net pressure between buy and sell orders at different levels of the limit order book and has since become a widely adopted feature for high-frequency prediction. Subsequent deep learning models, including DeepLOB [28], BINCTABL [22, 23], and TLOB [3], have demonstrated the value of using either raw order book features or engineered order flow features for mid-price forecasting. However, these models typically treat each asset independently and ignore the market-wide context in which assets influence each other.

While Cont et al. [8] explored cross-asset impacts of Order Flow Imbalance (OFI) through linear modeling, such approaches are inherently limited in capturing the nonlinear dependencies that characterize modern financial markets. Moreover, they do not leverage deep learning architectures capable of modeling complex interactions across assets, time, and order book structure.

Motivated by these limitations, we propose that mid-price forecasting can be significantly enhanced by expanding the model's perspective beyond single-stock inputs. By incorporating OFI features from a carefully selected group of peer assets, our approach captures a broader view of market dynamics, where cross-asset information flow plays a pivotal role in price formation. This hypothesis is supported by empirical evidence of inter-asset relationships, as well as Granger causality analysis conducted during the training phase to identify relevant peer stocks.

In this paper, we introduce **OF-MATNet**, a Transformer-based architecture designed to model the intricate interplay between temporal, cross-asset, and order book level features. Our contributions are threefold:

- **Multi-Asset Forecasting Framework:** We extend traditional single-asset LOB forecasting by constructing a multi-asset input representation, enabling the model to leverage a richer set of market signals and capture broader market conditions that influence individual stock returns. To ensure the model focuses on informative peer assets, we apply a Granger causality analysis on training data to pre-select assets with statistically significant predictive influence on the target stock.
- **Multi-Axis Attention Architecture:** We design **OF-MATNet** as a Transformer-based model that simultaneously attends over time, assets, and order book levels. This multi-axis attention mechanism enables the extraction of dependencies across multiple dimensions of market data, enhancing the model's ability to understand complex order flow patterns.
- **Insightful Empirical Analysis:** We evaluate OF-MATNet on 110 Nasdaq-listed assets and conduct comprehensive analyses to interpret the model's behavior. Our experiments reveal how the importance of cross-asset information grows with forecasting horizon, and we demonstrate a strong correlation between an asset's sensitivity to peer flows and its rate of forecasting improvement. These findings provide valuable insights for future research on multi-asset interactions in high-frequency market modeling.

These contributions highlight the practical and methodological benefits of modeling high-frequency market dynamics through nonlinear, causality-aware, multi-asset learning frameworks. The remainder of this paper is structured as follows: Section 2 reviews related work on mid-price forecasting and multi-asset modeling. Section 3 introduces key concepts, including order flow imbalance, Granger causality, and Transformers. Section 4 describes the OF-MATNet architecture and data. Sections 5 and 6 present experiments and analyses, covering causality results, forecasting performance, feature importance, horizon effects, and practical implications. Finally, Section 7 concludes with key findings and future directions.

2 Related Work

Mid-price prediction, and more generally, price forecasting, has been a longstanding topic in financial research. The limit order book (LOB), which provides detailed information about market depth and order flow, is a powerful source of features for short-term forecasting. Order messages drive LOB dynamics, making them essential for modeling market microstructure. Numerous studies have leveraged LOB data for price forecasting, primarily focusing on single-stock setups [14, 17, 26, 28, 29]. With the advent of deep learning and increased access to high-resolution LOB datasets, many recent works have adopted neural networks to capture temporal and spatial patterns in order book movements [13, 19, 20, 24].

One of the earliest and most influential deep learning models for LOB data is DeepLOB [28], which uses CNNs to extract spatial features and LSTMs to capture sequential dependencies. This architecture inspired many follow-up works. For example, Wallbridge et al. [26] replaced the LSTM module with a Transformer [25], while Lucchese et al. [14] proposed improvements to the input encoding and prediction targets to better support multi-horizon forecasting. Kolm et al. [12] incorporated order flow imbalance (OFI), a feature originally proposed by Cont et al. [9], and demonstrated its strong predictive power compared to raw book states.

OFI has emerged as a key feature in modern LOB forecasting. By summarizing the net buy/sell pressure, it has been shown to be highly correlated with short-term returns [9]. Zhang et al. [28] integrated OFI into the DeepLOB framework to form DeepOFI, showing its effectiveness over raw volume and price features. However, almost all these approaches focus on individual asset modeling, ignoring the fact that financial markets are inherently interconnected. Movements in one instrument often affect others through lead-lag effects, sectoral dynamics, or systematic flows.

To address this, Cont et al. [8] introduced a cross-impact framework using OFI from multiple stocks to forecast returns in a linear setting. While their results confirmed the presence of inter-asset predictive signals, their model was limited to linear dependencies and relatively coarse (one-minute) sampling, making it less suitable for high-frequency strategies. Other works have explored the decomposition of OFI by order type [21], or lead-lag effects among stocks for trading signal generation [2, 5]. Still, these works either remain linear or use handcrafted methods without leveraging deep learning.

Recent studies highlight the advantages of attention-based neural networks—particularly Transformers—for modeling financial time series. First introduced by Vaswani et al. [25], Transformers have been adapted to LOB data to capture temporal dependencies more flexibly than RNNs or CNNs. The TLOB model [3] demonstrated that dual attention mechanisms across time and LOB levels improve predictive accuracy, while Prata et al. [18] benchmarked various LOB forecasting models and concluded that attention-based architectures are more robust and accurate, especially in noisy high-frequency environments. Arroyo et al. [1] employed a Conv-Transformer to estimate order fill probabilities, emphasizing attention's role in enhancing both performance and interoperability. HLOB [4] enhances structural modeling by learning topological patterns in LOB depth.

Additionally, graph Neural Networks (GNNs) have been explored to model cross-asset dependencies by capturing market structure through asset correlations or return-based networks [6, 7, 16, 27]. However, many of these approaches rely on static graph structures that do not adapt to changing market conditions. While dynamic graph learning frameworks like DySTAGE [11] have been proposed, such methods often introduce additional complexity and computational overhead.

Our work contributes to this evolving literature by introducing a Transformer-based model—OF-MATNet—that incorporates OFI signals from a selected group of peer assets. We perform a rolling-window Granger causality analysis on the training data to identify influential peer assets for each target, and use this selection throughout model training and inference. This pre-selection strategy ensures scalability and interpretability, while effectively capturing cross-asset nonlinear dependencies through multi-axis attention mechanisms. We demonstrate that OF-MATNet consistently outperforms single-asset baselines in predictive accuracy, offering a practical and computationally efficient solution for high-frequency mid-price forecasting.

3 Background

This section provides the foundational concepts and methodologies used in our study. We first explain the structure and function of the Limit Order Book (LOB) and the role of order messages in updating it. Next, we define the mid-price and discuss its importance in financial trading. We then delve into the concept of Order Flow Imbalance (OFI), detailing its calculation and significance. Additionally, we introduce the Granger causality test and explain the core components of Transformer-based architectures, including positional encoding.

3.1 Limit Order Book and Order Messages

The Limit Order Book (LOB) is crucial in electronic trading systems, particularly in equity markets like Nasdaq. It provides a real-time record of outstanding buy and sell limit orders, organized by price level. The LOB is divided into two main components:

- **Bid side**: Contains all active buy orders.
- **Ask side**: Contains all active sell orders.

Order messages, which include information on order submissions, cancellations, and executions, enable the continuous update and reconstruction of the LOB. These messages can be recorded at very high frequencies, such as nanoseconds, though our dataset provides data at a millisecond resolution. This granularity offers substantial insight into market dynamics and participant behavior.

The mid-price is a key financial metric used in trading, calculated as the average of the best ask and best bid prices. This measure provides a stable estimate of the market price, reflecting the immediate balance between supply and demand. It is particularly useful in high-frequency trading (HFT) and market making, where accurate predictions of future price movements are critical for profitability.

3.2 Order Flow Imbalance (OFI)

Order Flow Imbalance (OFI) was introduced by Cont et al. [9] as a measure of the net order flow at the best bid and ask levels. It reflects the pressure on the market price due to buying and selling activities. For an order book event with index n and order book level ℓ, the ask order flow (aOF_n^ℓ) and bid order flow (bOF_n^ℓ) are defined as:

$$aOF_n^\ell = \begin{cases} -q_{a,n}^\ell, & \text{if } a_n^\ell > a_{n-1}^\ell \\ q_{a,n}^\ell - q_{a,n-1}^\ell, & \text{if } a_n^\ell = a_{n-1}^\ell \\ q_{a,n}^\ell, & \text{otherwise} \end{cases} \tag{1}$$

$$bOF_n^\ell = \begin{cases} q_{b,n}^\ell, & \text{if } b_n^\ell > b_{n-1}^\ell \\ q_{b,n}^\ell - q_{b,n-1}^\ell, & \text{if } b_n^\ell = b_{n-1}^\ell \\ -q_{b,n}^\ell, & \text{otherwise} \end{cases} \tag{2}$$

where $q_{a,n}^\ell$ and $q_{b,n}^\ell$ are the ask and bid sizes at level ℓ and time n, respectively. The cumulative OFI over a time interval $(t-h, t)$ is then given by:

$$OFI_t^h = \sum_{n=N(t-h)+1}^{N(t)} (bOF^n - aOF^n) \tag{3}$$

where $N(t)$ is the index of the last event occurring no later than time t.

3.3 Granger Causality Test

The Granger causality test is a statistical method used to determine whether one time series provides useful information for forecasting another. In financial markets, it helps assess whether the historical values of one variable, such as the Order Flow Imbalance (OFI) of a stock, can predict the future returns of another asset. This methodology is widely used to identify lead-lag relationships among stocks, which are essential for understanding cross-asset information flow in interconnected markets.

3.4 Transformer Encoder and Positional Encoding

The Transformer architecture, originally introduced by Vaswani et al. [25], has become a foundational model for sequential data processing due to its self-attention mechanism. Unlike recurrent networks, Transformers allow the model to attend to relevant parts of the input sequence regardless of their position, making them highly effective in capturing both local and long-range dependencies. Since Transformers do not inherently process sequence order, positional encodings are incorporated to inject information about the relative or absolute position of each element in the sequence.

4 Methodology

Our study aims to integrate two essential perspectives in financial market forecasting: the significance of cross-asset order flow dynamics and the power of deep learning architectures in modeling complex market patterns. Prior work by Cont et al. [8] has demonstrated that OFI across multiple stocks carries valuable predictive information for future returns. However, their approach was restricted to linear modeling frameworks and did not exploit the capabilities of deep learning models in capturing nonlinear dependencies and hierarchical structures inherent in LOB data.

A fundamental challenge in multi-asset modeling is the asynchronous nature of order book updates across different instruments. Order-event sequences vary significantly between assets, complicating the construction of consistent cross-asset feature representations. To address this, we adopt a timestamp-based synchronization approach, aggregating data into fixed 2-second intervals. This alignment not only mitigates the Epps effect but also facilitates coherent multi-asset feature representation and aligns with practical trading strategies that operate on time-based decision windows.

Our objective is to develop a forecasting framework that leverages OFI features from a causally relevant group of peer assets to predict the next-step return of a target stock. By incorporating multi-asset information within a deep learning framework, we aim to capture broader market conditions and inter-asset dependencies that influence individual stock returns, thereby enhancing predictive accuracy and robustness in high-frequency trading environments.

4.1 Research Problem Formulation

We formalize the task of multi-asset mid-price return forecasting as a supervised learning problem. Given historical OFI features from a group of N causally related assets over a fixed historical window of T time steps, the goal is to predict the next-step return of a target asset.

Let Δ denote the fixed timeframe interval used for data aggregation (e.g., 2 seconds). For each asset A_i in the selected set of peer assets $\mathcal{A} = \{A_1, A_2, ..., A_N\}$, we compute an OFI matrix $\text{OFI}^{(i)}_{t-T\Delta:t}$ representing order flow imbalance values across the top L order book levels over the past T time steps:

- $\text{OFI}^{(i)}_{t-T\Delta:t} \in \mathbb{R}^{T\times L}$ is the OFI sequence for asset A_i from time $t - T\Delta$ to t.
- T is the historical window length (sequence length).
- L is the number of order book levels considered (e.g., $L = 10$).
- N is the number of selected peer assets.

The mid-price P_t at timestamp t is defined as the average of the best bid (P_t^{bid}) and best ask (P_t^{ask}) prices:

$$P_t = \frac{P_t^{bid} + P_t^{ask}}{2} \tag{4}$$

For each timeframe Δ, we compute the average mid-price $\overline{P}_t$ as the mean of all mid-price values within that interval. The return over the next timeframe, $r_{t+\Delta}$, is defined as the relative change between the average mid-price of the next step and the current step:

$$r_{t+\Delta} = \frac{\overline{P}_{t+\Delta} - \overline{P}_t}{\overline{P}_t} \tag{5}$$

Our objective is to learn a function f_θ, parameterized by a neural network, that maps the historical multi-asset OFI sequence to the predicted next-step return of the target asset:

$$\hat{r}^{(\text{target})}_{t+\Delta} = f_\theta\left(\{\text{OFI}^{(i)}_{t-T\Delta:t}\}_{i=1}^{N}\right) \tag{6}$$

The problem requires modeling dependencies across three key dimensions:

- Temporal evolution of order flow within each asset.
- Cross-asset relationships among causally selected peer assets.
- Hierarchical structure of order book levels.

Our proposed model, **OF-MATNet**, is designed to effectively capture these dependencies using a multi-axis attention mechanism.

4.2 Data Description

The dataset used in this study is sourced from AlgoSeek, specifically the "US Equity Full Depth for All Symbols from NASDAQ ITCH Historical Data." This comprehensive dataset includes order messages consolidated and ordered by millisecond timestamps, enabling the reconstruction of the full depth of the LOB at any moment during the trading day. The data spans the entire trading session, including early and late trading hours from 04:00 to 20:00 EST, providing a detailed view of market activities throughout the day. [1]

We utilize a 33-day period of Nasdaq order book data encompassing various market conditions, with dataset partitioning details provided in Section 5.1.

We select a diverse pool of 110 assets, including stocks, ETFs, and indices, spanning multiple sectors such as technology, finance, healthcare, energy, consumer goods, and industrials. The asset list includes high-liquidity blue-chip stocks (e.g., AAPL, MSFT, NVDA), sector-specific ETFs (e.g., XLE, XLK, XLY), broad-market indices (e.g., SPY, QQQ, DIA), and assets from sectors like energy (e.g., CVX, OXY, SLB), consumer discretionary (e.g., AMZN, TSLA), and financials (e.g., JPM, BAC, GS). This diversity ensures that our dataset reflects a wide range of market behaviors and interaction patterns, providing a robust foundation for multi-asset modeling.

For each asset, we reconstruct the LOB using millisecond-level order messages and compute the OFI across the top $L = 10$ price levels. To synchronize data across assets, we aggregate events into fixed 2-second intervals, ensuring that the input features for all assets are temporally aligned.

Each input sample is constructed as a tensor of shape $[B, T, N, L]$, where B is the batch size, $T = 40$ is the historical sequence length , $N = 50$ is the number of selected assets per target (chosen via Granger causality), and $L = 10$ is the number of order book levels. This synchronized and diversified dataset forms the foundation for training and evaluating our proposed OF-MATNet model.

4.3 Model Architecture

Our proposed model, **OF-MATNet**, is a Transformer-based architecture designed to capture complex dependencies across temporal sequences, cross-asset relationships, and order book level hierarchies. The architecture is composed of three specialized attention paths—Temporal, Cross-Asset, and Level—that operate in parallel to extract representations from multi-dimensional order flow data.

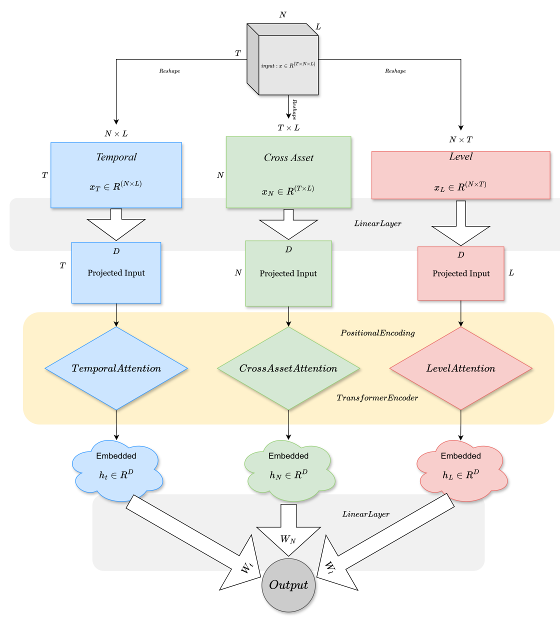

Figure 1: OF-MATNet Architecture: Multi-axis attention paths attend over temporal, cross-asset, and order flow level dimensions.

Input Representation and Reshaping: The input tensor has the shape $[B, T, N, L]$, where B is the batch size, $T = 40$ is the sequence length, $N = 50$ is the number of selected assets, and $L = 10$ represents the order book levels. To enable axis-specific attention, the input is reshaped along three dimensions:

[1] We extend our gratitude to AlgoSeek for providing this dataset, which has been instrumental in enabling our analysis.

- Temporal Path: reshaped to $[B, T, N \times L]$.
- Cross-Asset Path: reshaped to $[B, N, T \times L]$.
- Level Path: reshaped to $[B, L, T \times N]$.

Projection and Positional Encoding: Each path is projected to a shared embedding dimension D through separate linear layers. Temporal and Level paths receive positional encodings to preserve sequential and hierarchical information, respectively. Cross-Asset path embeddings are distinguished inherently by asset identity.

Transformer Encoder Modules: Each axis-specific sequence is processed through an independent Transformer Encoder. The attention mechanisms within these encoders capture:

- Temporal dependencies within each asset-level pair.
- Cross-asset influences through interactions among selected peer assets.
- Level-wise contributions across the order book depth.

Embedding Fusion and Output Layer: The outputs from Temporal (h_T), Cross-Asset (h_N), and Level (h_L) paths are concatenated to form a unified representation:

$$h_{final} = \text{Concat}(h_T, h_N, h_L) \tag{7}$$

A final linear layer maps $h_{final} \in \mathbb{R}^{3D}$ to a scalar output predicting the next-step return of the target asset:

$$\hat{r}_{t+\Delta}^{(\text{target})} = W_{out} \cdot h_{final} \tag{8}$$

The model is trained to minimize Mean Squared Error (MSE) loss, optimizing its ability to forecast short-term returns using causally informative multi-asset order flow data.

5 Experimental Results and Discussion

In this section, we present the experimental setup, evaluation methodology, and empirical results of our proposed **OF-MATNet** model. We first outline the dataset partitioning, model configuration, and evaluation metrics, followed by a summary of the Granger causality analysis used for peer asset selection. We then compare the forecasting accuracy of **OF-MATNet** against its single-asset variant (**OF-SATNet**) and state-of-the-art baselines, including DeepLOB, BINCTABL, and TLOB, to demonstrate the advantages of incorporating multi-asset order flow information. Additionally, we provide interpretability analyses to investigate the contributions of temporal, cross-asset, and level-wise features, and explore how factors such as forecasting horizon influence model performance.

5.1 Experimental Setup

We conduct our experiments on Nasdaq Limit Order Book (LOB) data covering a 33-day period from June 9, 2022, to July 28, 2022, encompassing a range of market conditions. The dataset is divided into three subsets:

- **Training Set:** First 21 trading days.
- **Validation Set:** Following 4 trading days.
- **Test Set:** Remaining 8 trading days.

For each asset, we reconstruct the LOB using millisecond-level order messages and compute OFI across the top $L = 10$ order book levels. To synchronize data across assets, events are aggregated into fixed 2-second intervals, ensuring temporal alignment. Each input sample consists of $T = 40$ historical intervals, used to forecast the next mid-price return of the target asset.

Input tensors are constructed with shape $[B, T, N, L]$, where B is the batch size, $N = 50$ is the number of selected peer assets per target (chosen via Granger causality analysis), and $L = 10$ is the number of order book levels. In the single-asset variant (OF-SATNet), N is set to 1.

Model performance is evaluated using the coefficient of determination (R^2) on the test set, defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(r_i^{(\text{target})} - \hat{r}_i^{(\text{target})} \right)^2}{\sum_{i=1}^{n} \left(r_i^{(\text{target})} - \overline{r}^{(\text{target})} \right)^2} \tag{9}$$

where $r_i^{(\text{target})}$ denotes the true return of the target asset at sample i, $\hat{r}_i^{(\text{target})}$ is the predicted return, and $\overline{r}^{(\text{target})}$ is the mean of the true returns. An R^2 score closer to 1 indicates better predictive performance.

5.2 Granger Causality Analysis

To identify influential peer assets, we conduct a rolling-window Granger causality analysis on the training data. The test is applied in two stages: once on mid-price returns and once on aggregated OFI, computed as the sum of OFI across the top 10 order book levels. For each asset pair (A_i, A_j), we test whether the historical values of A_i Granger-cause the future returns or OFI of A_j, using a rolling window of 1,000 intervals with a step size of 600. A directed edge from A_i to A_j is drawn if causality is detected in more than 70% of windows for returns and more than 85% for OFI. Edge thickness is proportional to the frequency of detected causality.

Figure 2 shows the resulting causality networks. Both return-based and OFI-based graphs are densely connected, highlighting strong interdependencies among assets. We observe that smaller or less liquid stocks tend to receive more incoming edges, indicating they are influenced by larger assets. Conversely, leading assets, such as indices and high-liquidity assets, act as causal sources with more outgoing edges. These findings emphasize the importance of modeling cross-asset influences in forecasting tasks.

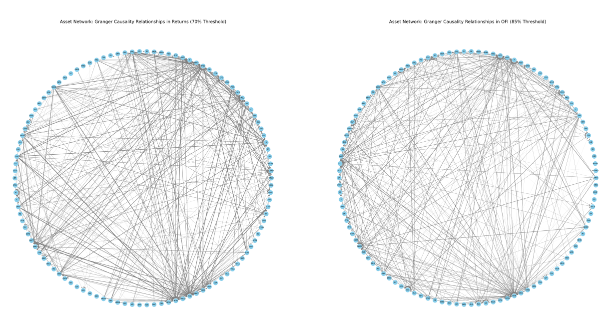

Figure 2: Network of Granger causality relationships among asset Returns (left) and OFIs (right).

5.3 Single-Asset Benchmarking

To establish a baseline for evaluating our architecture, we first assess the performance of the single-asset variant, **OF-SATNet**, in comparison with state-of-the-art models, including DeepLOB, BINCTABL, and TLOB. This experiment is designed to evaluate the effectiveness of our model in a traditional single-stock forecasting setup, without incorporating cross-asset information.

We select a diversified subset of 16 Nasdaq-listed assets, spanning various sectors and liquidity profiles. Table 1 reports the R^2 scores for each model on these assets. As shown, OF-SATNet achieves the highest R^2 in several cases, particularly on stocks such as AAPL, JPM, CVX, and DIS, demonstrating the effectiveness of our multi-axis attention mechanism in capturing temporal and order book level dynamics. However, in other cases, models like DeepLOB outperform, especially on assets where single-stock dynamics dominate, such as TSLA, META, and AMD.

Table 1: R^2 comparison of single-asset models. The best performing model per asset is highlighted in bold.

Asset	OF-SATNet	BINCTABL	DeepLOB	TLOB
SPY	0.202	0.192	**0.208**	0.187
XLE	**0.194**	0.174	0.193	0.170
AAPL	**0.238**	0.191	0.228	0.190
TSLA	0.179	0.171	**0.196**	0.173
JPM	**0.163**	0.109	0.155	0.116
JNJ	0.151	0.143	**0.170**	0.132
WMT	**0.139**	0.110	0.134	0.109
META	0.182	0.173	**0.204**	0.171
CVX	**0.190**	0.155	0.188	0.157
NEE	**0.137**	0.107	0.125	0.099
AMD	0.226	0.193	**0.237**	0.189
DIS	**0.162**	0.129	0.155	0.125
NFLX	0.122	0.115	**0.136**	0.114
BA	**0.088**	0.079	0.085	0.075
INTC	0.158	0.143	**0.174**	0.159
AMZN	0.203	0.160	**0.204**	0.157

These results indicate that while OF-SATNet is competitive with existing methods in the single-asset setting, its performance varies across different assets. This variation highlights the limitation of relying solely on single-stock inputs for forecasting. To address this, we propose extending the model's input space to include information from a causally relevant group of peer assets. By incorporating multi-asset order flow information, we aim to capture broader market dynamics that influence the target asset's price movements, thereby enhancing predictive accuracy. The results of this multi-asset forecasting approach are presented in the next section.

5.4 Multi-Asset Forecasting

Having established a baseline with the single-asset variant (**OF-SATNet**), we now evaluate the performance of the proposed **OF-MATNet** model, which incorporates causally selected peer assets for each target asset. Table 2 presents the R^2 comparison of OF-MATNet against state-of-the-art baseline models, including DeepLOB, BINCTABL, and TLOB. The results demonstrate that OF-MATNet consistently outperforms these baselines across the majority of assets. Statistically significant improvements ($p < 0.05$) are marked with an asterisk. For significance testing, we apply a paired t-test on the prediction errors, following standard practices for evaluating model improvements in time series forecasting [10].

Table 2: R^2 comparison of OF-MATNet vs baseline models with significance. * indicates statistically significant improvement ($p < 0.05$).

Asset	OF-MATNet	BINCTABL	DeepLOB	TLOB
SPY	**0.247**	0.192*	0.208*	0.187*
XLE	**0.201**	0.174*	0.193*	0.170*
AAPL	**0.254**	0.191*	0.228*	0.190*
TSLA	**0.242**	0.171*	0.196*	0.173*
JPM	**0.176**	0.109*	0.155*	0.116*
JNJ	**0.186**	0.143*	0.170*	0.132*
WMT	**0.157**	0.110*	0.134*	0.109*
META	**0.206**	0.173*	0.204	0.171*
CVX	0.147	0.155	**0.188**	0.157
NEE	**0.132**	0.107*	0.125*	0.099*
AMD	**0.253**	0.193*	0.237*	0.189*
DIS	**0.207**	0.129*	0.155*	0.125*
NFLX	**0.152**	0.115*	0.136*	0.114*
BA	**0.123**	0.079*	0.085*	0.075*
INTC	**0.178**	0.143*	0.174	0.159*
AMZN	**0.228**	0.160*	0.204*	0.157*

In addition to outperforming traditional baselines, OF-MATNet also demonstrates clear improvements over its single-asset counterpart (OF-SATNet). To quantify this, we evaluate both models across all 110 assets in our dataset. Figure 3 visualizes the R^2 improvement percentage of OF-MATNet over OF-SATNet for each asset. Bars are colored based on the statistical significance of the improvement, determined using a paired t-test at the 5% level.

Incorporating multi-asset order flow information enhances forecasting performance in most cases, with some assets improving by over 50%. Low-liquidity assets benefit the most by leveraging signals from dominant peers, while market leaders, acting as causal sources, show more modest gains. Additionally, assets like CVX and NEE sometimes perform similarly in single-asset models due to limited sector coverage in our peer asset pool and external macro factors—such as commodity prices and regulations—that influence their price dynamics beyond peer flows.

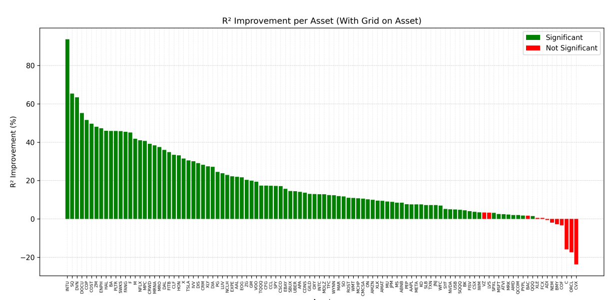

Figure 3: R^2 Improvement (%) of OF-MATNet over OF-SATNet per Asset.

Overall, OF-MATNet achieves statistically significant improvements over OF-SATNet in **89% of assets** (at a 5% significance level).

Furthermore, the R^2 score improves in more than **90% of assets**, highlighting the consistent advantage of incorporating multi-asset order flow information for forecasting accuracy. These results underscore the robustness and generalizability of our causality-aware multi-asset modeling approach across diverse asset classes and market conditions.

To better understand the factors contributing to these variations in improvement, we analyze the relationship between the strength of causality influence an asset receives from other assets and its corresponding R^2 improvement. Figure 4 illustrates this relationship, where the x-axis represents the average percentage of Granger causality influence from peer assets, and the y-axis denotes the R^2 improvement achieved by OF-MATNet over OF-SATNet.

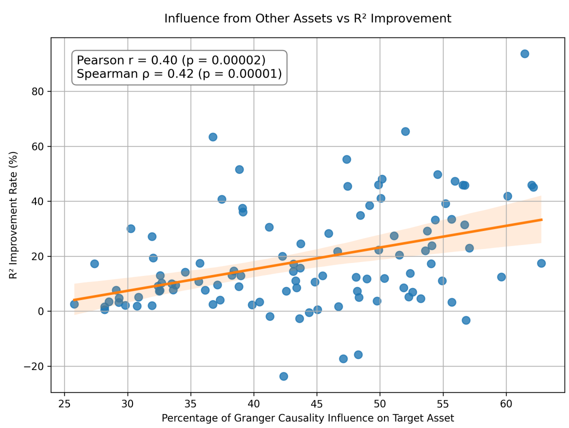

Figure 4: Relationship between Granger Causality Influence and R^2 Improvement Rate.

The analysis reveals a clear positive correlation, quantified by Pearson's $r = 0.40$ (p = 0.00002) and Spearman's $\rho = 0.42$ (p = 0.00001). This suggests that assets which are more influenced by peer flows benefit more from multi-asset modeling, as they can leverage additional predictive signals not captured in single-asset frameworks. These findings emphasize the importance of causality-aware asset selection and highlight the asymmetric nature of cross-asset dependencies in financial markets.

5.5 Feature Importance in OF-MATNet

We analyze feature importance from two perspectives: (1) peer asset contributions in forecasting, and (2) the relative influence of temporal, cross-asset, and level-wise features.

Our analysis reveals a consistent pattern where sector-specific ETFs and large-cap peers within the same sector rank among the top influential assets for each target. Table 3 summarizes key examples, illustrating how the model leverages both asset-specific signals and broader sector-wide dynamics. For instance, XLK (Technology ETF) and MSFT are major contributors in TSLA's prediction, while XLF (Financial ETF) and GS are prominent for JPM. Similarly, for JNJ, healthcare-focused ETF XLV and assets like PFE and UNH rank among the most influential.

We also evaluate the contribution of temporal, cross-asset, and level-wise features by analyzing the learned projection weights. As

Table 3: Top-5 Influential Peer Assets per Target.

Target	1st	2nd	3rd	4th	5th
TSLA	SQQQ	XLK	XLY	MSFT	CDNS
JPM	JPM	XLF	GS	BAC	SPY
JNJ	JNJ	XLV	PFE	UNH	XLF

shown in Table 4, temporal dependencies dominate, accounting for 86% of the model's output influence. However, cross-asset and level-wise signals still contribute meaningful information, enhancing predictive accuracy.

Table 4: Axis-wise Contribution Weights.

	Temporal	Asset	Level
Contribution	86%	8%	6%

5.6 Forecast Horizon Impact Analysis

To assess the effect of forecasting horizon on model performance, we evaluate OF-MATNet across multiple horizons (1, 2, 4, and 6 steps ahead). Table 5 reports the relative R^2 improvement of OF-MATNet over OF-SATNet across six representative assets. Results indicate that the benefit of incorporating multi-asset information becomes more pronounced as the forecasting horizon extends. For example, TSLA exhibits a 30% improvement at Horizon 1, which amplifies to over 129,000% at Horizon 6.

Table 5: R^2 Improvement (%) of OF-MATNet over OF-SATNet at Different Forecast Horizons.

Asset	Horizon 1	Horizon 2	Horizon 4	Horizon 6
XLE	0.51	0.47	25.24	151.05
AMZN	10.00	8.10	435614.62	71439.46
TSLA	30.54	39.63	3611.04	129978.61
PLTR	45.83	91.73	135.55	232.55
CVX	-23.73	18.73	79.89	109.98
NFLX	41.05	59.24	29433.17	9238.06

2

Interestingly, some assets that showed negligible improvement at short horizons, benefit significantly at longer horizons. For example, CVX's R^2 improvement grows from negative value at Horizon 1 to over 100% at Horizon 6, underscoring how cross-asset dependencies amplify with longer prediction intervals.

We further analyze how feature importance shifts across forecasting horizons for AMZN by visualizing aggregated SHAP values [15] over asset and level dimensions (Figure 5).

As the horizon extends, feature importance spreads downward to a broader set of assets and shifts rightward to deeper order book levels. This reflects the increasing reliance on cross-asset signals and deeper liquidity layers for longer-term predictions, as also mentioned in [8].

[2] *Note:* extremely large relative improvements occur when the baseline R^2 at long horizons is near zero; absolute gains remain small in those cases.

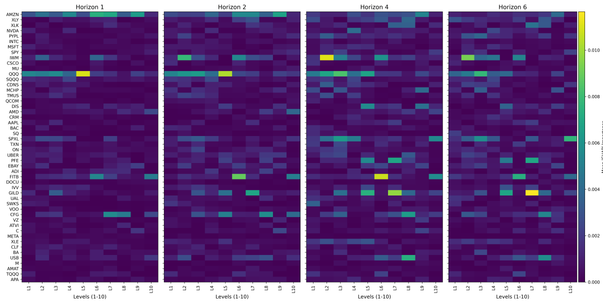

Figure 5: 2D Feature Importance across Horizons for AMZN (Assets vs Levels).

6 Discussion and Practical Implications

Our findings demonstrate that incorporating causally relevant peer assets into forecasting models enhances predictive accuracy, particularly at longer horizons where single-asset signals diminish. The consistent importance of sector ETFs and large-cap peers highlights the critical role of sector-level dynamics in price formation. From a practical standpoint, while adding more assets improves model performance, the marginal benefits decrease as additional assets are included. This underscores the necessity of informed asset selection strategies to maintain an effective balance between model complexity and predictive gain. These insights are particularly valuable for high-frequency trading and market-making applications, where leveraging cross-asset information flow can provide a competitive edge in price prediction and execution strategies.

For latency-sensitive applications, we also experimented with a lightweight variant of OF-MATNet, using only temporal attention and a single aggregated OFI feature per asset. This simplified model maintained strong performance, outperforming OF-SATNet in over 75% of assets, while significantly reducing computational overhead.

7 Conclusion

In this paper, we proposed OF-MATNet, a causality-aware multi-asset forecasting framework that leverages order flow imbalance features from a selected group of peer assets to enhance mid-price prediction accuracy. Through extensive experiments on Nasdaq order book data, we demonstrated that incorporating multi-asset information significantly improves forecasting performance over single-asset models and established baselines, especially at longer prediction horizons.

Our analyses revealed that sectoral relationships and market-wide dynamics play a pivotal role in forecasting improvements, as evidenced by the importance of sector ETFs and large-cap peers in the model's feature attributions. Furthermore, we found that the magnitude of improvement is strongly correlated with factors such as an asset's causality strength from peers and the forecasting horizon length. While temporal dependencies dominate the model's decision process, cross-asset and order book structure information contribute meaningfully to refining predictions.

For future research, extending this framework with Graph Neural Networks (GNNs) that dynamically learn cross-asset relationships, such as attention-based GATs, could further enhance the model's adaptability to evolving market structures.

Acknowledgments

We thank **AlgoSeek** for providing access to the US Equity Full Depth (NASDAQ ITCH) dataset and technical support on its usage.

References

[1] Álvaro Arroyo, Álvaro Cartea, Francisco Moreno-Pino, and Stefan Zohren. 2024. Deep Attentive Survival Analysis in Limit Order Books: Estimating Fill Probabilities with Conv-Transformers. *Quantitative Finance* (2024).

[2] Stefanos Bennett, Mihai Cucuringu, and Gesine Reinert. 2022. Lead–lag detection and network clustering for multivariate time series with an application to the US equity market. *Machine Learning* 111, 12 (2022), 4497–4538.

[3] Luca Berti and Gjergji Kasneci. 2025. TLOB: A Transformer Model with Dual Attention for Price Trend Prediction with LOB Data. *arXiv preprint arXiv:2502.15757* (2025).

[4] Andrea Briola and Tomaso Aste. 2024. HLOB: A Deep Learning Model for Structured Limit Order Book Forecasting. *Expert Systems with Applications* 239 (2024), 122116.

[5] Álvaro Cartea, Mihai Cucuringu, and Qi Jin. 2023. Detecting Lead-Lag Relationships in Stock Returns and Portfolio Strategies. *Available at SSRN* (2023).

[6] Qinkai Chen and Christian-Yann Robert. 2022. Multivariate realized volatility forecasting with graph neural network. In *Proceedings of the third acm international conference on ai in finance*. 156–164.

[7] Yingmei Chen, Zhongyu Wei, and Xuanjing Huang. 2018. Incorporating corporation relationship via graph convolutional neural networks for stock price prediction. In *Proceedings of the 27th ACM international conference on information and knowledge management*. 1655–1658.

[8] Rama Cont, Mihai Cucuringu, and Chao Zhang. 2023. Cross-impact of order flow imbalance in equity markets. *Quantitative Finance* 23, 10 (2023), 1373–1393.

[9] Rama Cont, Arseniy Kukanov, and Sasha Stoikov. 2014. The price impact of order book events. *Journal of financial econometrics* 12, 1 (2014), 47–88.

[10] Francis X Diebold and Roberto S Mariano. 1995. Comparing predictive accuracy. *Journal of Business & Economic Statistics* 13, 3 (1995), 253–263.

[11] Jingyi Gu, Junyi Ye, Ajim Uddin, and Guiling Wang. 2024. DySTAGE: Dynamic Graph Representation Learning for Asset Pricing via Spatio-Temporal Attention and Graph Encodings. In *Proceedings of the 5th ACM International Conference on AI in Finance*. 388–396.

[12] Petter N Kolm, Jeremy Turiel, and Nicholas Westray. 2023. Deep order flow imbalance: Extracting alpha at multiple horizons from the limit order book. *Mathematical Finance* 33, 4 (2023), 1044–1081.

[13] Shijie Lai, Mingxian Wang, Shengjie Zhao, and Gonzalo R Arce. 2023. Predicting high-frequency stock movement with differential transformer neural network. *Electronics* 12, 13 (2023), 2943.

[14] Lorenzo Lucchese, Mikko S Pakkanen, and Almut ED Veraart. 2024. The short-term predictability of returns in order book markets: a deep learning perspective. *International Journal of Forecasting* (2024).

[15] Scott M Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*, Vol. 30. 4765–4774.

[16] Daiki Matsunaga, Toyotaro Suzumura, and Toshihiro Takahashi. 2019. Exploring graph neural networks for stock market predictions with rolling window analysis. *arXiv preprint arXiv:1909.10660* (2019).

[17] Adamantios Ntakaris, Martin Magris, Juho Kanniainen, Moncef Gabbouj, and Alexandros Iosifidis. 2018. Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods. *Journal of Forecasting* 37, 8 (2018), 852–866.

[18] Matteo Prata, Giuseppe Masi, Leonardo Berti, Viviana Arrigoni, Andrea Coletta, Irene Cannistraci, Svitlana Vyetrenko, Paola Velardi, and Novella Bartolini. 2024. Lob-based deep learning models for stock price trend prediction: a benchmark study. *Artificial Intelligence Review* 57, 5 (2024), 116.

[19] Mostafa Shabani, Dat Thanh Tran, Martin Magris, Juho Kanniainen, and Alexandros Iosifidis. 2022. Multi-head temporal attention-augmented bilinear network for financial time series prediction. In *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 1487–1491.

[20] Justin Sirignano and Rama Cont. 2021. Universal features of price formation in financial markets: perspectives from deep learning. In *Machine learning and AI in finance*. Routledge, 5–15.

[21] Bogdan Sitaru, Anisoara Calinescu, and Mihai Cucuringu. 2023. Order flow decomposition for price impact analysis in equity limit order books. In *Proceedings of the Fourth ACM International Conference on AI in Finance*. 637–645.

[22] Dat Thanh Tran, Alexandros Iosifidis, Juho Kanniainen, and Moncef Gabbouj. 2018. Temporal attention-augmented bilinear network for financial time-series

data analysis. *IEEE transactions on neural networks and learning systems* 30, 5 (2018), 1407–1418.
[23] Dat Thanh Tran, Juho Kanniainen, Moncef Gabbouj, and Alexandros Iosifidis. 2021. Data normalization for bilinear structures in high-frequency financial time-series. In *2020 25th International conference on pattern recognition (ICPR)*. IEEE, 7287–7292.
[24] Avraam Tsantekidis, Nikolaos Passalis, Anastasios Tefas, Juho Kanniainen, Moncef Gabbouj, and Alexandros Iosifidis. 2017. Forecasting stock prices from the limit order book using convolutional neural networks. In *2017 IEEE 19th conference on business informatics (CBI)*, Vol. 1. IEEE, 7–12.
[25] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* 30 (2017).
[26] James Wallbridge. 2020. Transformers for limit order books. *arXiv preprint arXiv:2003.00130* (2020).
[27] Tao Yin, Chenzhengyi Liu, Fangyu Ding, Ziming Feng, Bo Yuan, and Ning Zhang. 2022. Graph-based stock correlation and prediction for high-frequency trading systems. *Pattern Recognition* 122 (2022), 108209.
[28] Zihao Zhang, Stefan Zohren, and Stephen Roberts. 2019. Deeplob: Deep convolutional neural networks for limit order books. *IEEE Transactions on Signal Processing* 67, 11 (2019), 3001–3012.
[29] Ban Zheng, Eric Moulines, and Frédéric Abergel. 2012. Price jump prediction in limit order book. *arXiv preprint arXiv:1204.1381* (2012).