VOL. 3, NO. 10

OMI Research Newsletter – August 2024

The Oxford-Man Institute of Quantitative Finance (OMI) is a world-leading centre for interdisciplinary research in financial markets. The OMI addresses fundamental problems in quantitative finance with a strong focus on machine learning and data driven models.

*The **OMI Research Newsletter** is a curated summary that covers the most recent research, insights, and tools related to financial decision making.*

Highlights

Research focus. Key contributions in this month's edition include works from the OMI, University of Cambridge, Queen Mary University, Northwestern University, and University of Toronto. Cartea, Jaimungal and Sánchez-Betancourt (2024) study equilibrium in a broker-mediated market. Moreno-Pino et al. (2024) introduce a variation of the Transformer model for continuous-time sequences which incurs significantly lower computational costs. Yang and Shestopaloff (2024) employ Bayesian methods for high dimensional cointegration. Finally, Chopra (2024) shows that without commitment, costly signaling, or intermediaries, markets fail when there is private information.

Machine Learning & Statistics

Rough Transformers: Lightweight Continuous-Time Sequence Modelling with Path Signatures

F. MORENO-PINO, Á. ARROYO, H. WALDON, X. DONG, Á. CARTEA

OXFORD-MAN INSTITUTE

Time-series data in real-world settings typically exhibit long-range dependencies and are observed at non-uniform intervals. In these settings, traditional sequence-based recurrent models struggle. To overcome this, researchers often replace recurrent architectures with Neural ODE-based models to account for irregularly sampled data and use Transformer-based architectures to account for long-range dependencies. Despite the success of these two approaches, both incur very high computational costs for input sequences of even moderate length. To address this challenge, we introduce the Rough Transformer, a variation of the Transformer model that operates on continuous-time representations of input sequences and incurs significantly lower computational costs. In particular, we propose *multi-view signature attention*, which uses path signatures to augment vanilla attention and to capture both local and global (multi-scale) dependencies in the input data, while remaining robust to changes in the sequence length and sampling frequency and yielding improved spatial processing. We find that, on a variety of time-series-related tasks, Rough Transformers consistently outperform their vanilla attention counterparts while obtaining the representational benefits of Neural ODE-based models, all at a fraction of the computational time and memory resources.

Learning Low-dimensional Latent Dynamics from High-dimensional Observations: Non-asymptotics and Lower Bounds

Y. ZHANG, S. TALEBI, N. LI

HARVARD UNIVERSITY

In this paper, we focus on learning a linear time-invariant (LTI) model with low-dimensional latent variables but high-dimensional observations. We provide an algorithm that recovers the high-dimensional features, i.e. column space of the observer, embeds the data into low dimensions and learns the low-dimensional model parameters. Our algorithm enjoys a sample complexity guarantee of order $\tilde{\mathcal{O}}(n/\epsilon^2)$, where n is the observation dimension. We further establish a fundamental lower bound indicating this complexity bound is optimal up to logarithmic factors and dimension-independent constants. We show that this inevitable linear factor of n

is due to the learning error of the observer's column space in the presence of high-dimensional noises. Extending our results, we consider a meta-learning problem inspired by various real-world applications, where the observer column space can be collectively learned from datasets of multiple LTI systems. An end-to-end algorithm is then proposed, facilitating learning LTI systems from a meta-dataset which breaks the sample complexity lower bound in certain scenarios.

Low Volatility Stock Portfolio Through High Dimensional Bayesian Cointegration

P. R. Yang, A. Y. Shestopaloff
University of Cambridge, Queen Mary University of London

We employ a Bayesian modelling technique for high dimensional cointegration estimation to construct low volatility portfolios from a large number of stocks. The proposed Bayesian framework effectively identifies sparse and important cointegration relationships amongst large baskets of stocks across various asset spaces, resulting in portfolios with reduced volatility. Such cointegration relationships persist well over the out-of-sample testing time, providing practical benefits in portfolio construction and optimization. Further studies on drawdown and volatility minimization also highlight the benefits of including cointegrated portfolios as risk management instruments.

Challenges in Training PINNs: A Loss Landscape Perspective

P. Rathore, W. Lei, Z. Frangella, L. Lu, M. Udell
Stanford University

This paper explores challenges in training Physics-Informed Neural Networks (PINNs), emphasizing the role of the loss landscape in the training process. We examine difficulties in minimizing the PINN loss function, particularly due to ill-conditioning caused by differential operators in the residual term. We compare gradient-based optimizers Adam, L-BFGS, and their combination Adam+L-BFGS, showing the superiority of Adam+L-BFGS, and introduce a novel second-order optimizer, NysNewton-CG (NNCG), which significantly improves PINN performance. Theoretically, our work elucidates the connection between ill-conditioned differential operators and ill-conditioning in the PINN loss and shows the benefits of combining first- and second-order optimization methods. Our work presents valuable insights and more powerful optimization strategies for training PINNs, which could improve the utility of PINNs for solving difficult partial differential equations.

Position: Embracing Negative Results in Machine Learning

F. Karl, L. M. Kemeter, G. Dax, P. Sierak
Fraunhofer Institute, Ludwig-Maximilians-Universitat Munchen

Publications proposing novel machine learning methods are often primarily rated by exhibited predictive performance on selected problems. In this position paper we argue that predictive performance alone is not a good indicator for the worth of a publication. Using it as such even fosters problems like inefficiencies of the machine learning research community as a whole and setting wrong incentives for researchers. We therefore put out a call for the publication of "negative" results, which can help alleviate some of these problems and improve the scientific output of the machine learning research community. To substantiate our position, we present the advantages of publishing negative results and provide concrete measures for the community to move towards a paradigm where their publication is normalized.

Financial Statement Analysis with Large Language Models

A. Kim, M. Muhn, V. Nikolaev
The University of Chicago

We investigate whether an LLM can successfully perform financial statement analysis in a way similar to a professional human analyst. We provide standardized and anonymous financial statements to GPT4 and instruct the model to analyze them to determine the direction of future earnings. Even without any narrative or industry-specific information, the LLM outperforms financial analysts in its ability to predict earnings changes. The LLM exhibits a relative advantage over human analysts in situations when the analysts tend to struggle. Furthermore, we find that the prediction accuracy of the LLM is on par with the performance of a narrowly trained state-of-the-art ML model. LLM prediction does not stem from its training memory. Instead, we find that the LLM generates useful narrative insights about a company's future performance. Lastly, our trading strategies based on GPT's predictions yield a higher Sharpe ratio and alphas than strategies based on other models. Taken together, our results suggest that LLMs may take a central role in decision-making.

Finance and Economics

Cohesion, Ideology, and Tolerance

P. Allmis
University of Cambridge

Agents with different ideologies often form alliances to achieve their goals. Paradoxically, ideologically similar

agents are often opponents. In this paper, ideologically heterogeneous agents choose the ideological composition of their neighborhood, their tolerance, and invest into connections. The resulting weighted network describes allies, opponents, and strengths. Disputes with opponents determine benefits, which increase in an agent's strength and cohesion. Cohesive agents have fewer mutual allies with opponents. In equilibrium, the network is segregated when cohesion is effective enough and some agents tolerate ideologically distant types to oppose closer ones. Subsidizing connections dampens polarization in societies on the verge of segregation.

On the Expansion of Risk Pooling

R. Feng, S. Kim, M. Anthropelos

University of Piraeus, Tsinghua University, University of Illinois at Urbana-Champaign

Risk pooling has become an increasingly critical tool for managing risks among corporations, institutions, states, and nations, with examples including multinational pooling, decentralized insurance schemes, catastrophe risk pooling, burden sharing for nuclear accidents. While there has been a rich literature on such practices, little is known from a theoretical viewpoint regarding the operational strategies of risk pools and in particular on issues about the effect of a pool's expansion on each existing member's welfare. This paper is the first to explore these issues by establishing different notions of consensus for the expansion of a risk pool: strong consensus, where both existing members and new candidates improve their risk measures due to the pool's expansion, and weak consensus, which refers to the willingness of existing members to remain in the pool. Under optimal risk sharing for each pool, we show that its properties regarding expansion's effects depend strongly on the underlying pricing rule. For instance, only with a risk-adjusted equilibrium pricing are existing members willing to accept highly risky new members, whereas simple linear pricing excludes such members from the pool. Additionally, the impact of exogenous reinsurance on consensus is analyzed under both pricing rules.

Getting the Agent to Wait

M. Saeedi, Y. Shen, A. Shourideh

Carnegie Mellon University

We examine the strategic interaction between an expert (principal) maximizing engagement and an agent seeking swift information. Our analysis reveals: When priors align, relative patience determines optimal disclosure – impatient agents induce gradual revelation, while impatient principals cause delayed, abrupt revelation. When priors disagree, catering to the bias often emerges, with the principal initially providing signals aligned with the agent's bias. With private agent beliefs, we observe two phases: one engaging both agents, followed by catering to one type. Comparing personalized and non-personalized strategies, we find faster information revelation in the non-personalized case, but higher quality information in the personalized case.

Certifying Lemons

H. Chopra

Northwestern University

This paper examines an adverse selection environment where a sender with private information (high or low ability) tries to convince a receiver of their high ability. Without commitment or costly signaling, market failure can occur. Certification intermediaries reduce these frictions by enabling signaling through hard information. This paper focuses on a monopolistic certifier and its impact on equilibrium welfare and certificate design. Key findings show that the certifier often provides minimal information, pooling senders of varying abilities and leaving low rents for high-type senders, which typically disadvantages the receiver. However, when precise information is demanded, the certifier screens the sender perfectly, benefiting the receiver. Thus, the monopolistic intermediary has an ambiguous effect on market efficiency. The results emphasize the importance of high certification standards, which drive low-ability senders out of the market. Conditions for such equilibria are characterized, showing how simple threshold strategies by the receiver induce first-best outcomes. Additionally, the relationship between the characteristics of offered certificates and welfare is identified.

What does a dynamic oligopoly maximize? The continuous time Markov case

J. P. Rincón-Zapatero

Universidad Carlos III de Madrid

We analyze the question of whether the outcome of an oligopoly exploiting a nonrenewable resource can be replicated by a related monopoly, within the framework of continuous time and Markov Perfect Nash Equilibrium. We establish necessary and sufficient conditions and find explicit solutions in some cases. Also, very simple models with externalities are shown which Nash equilibrium cannot be replicated in a monopoly.

Mathematical Finance

Nash Equilibrium between Brokers and Traders

Á. Cartea, S. Jaimungal, L. Sánchez-Betancourt

Oxford-Man Institute, University of Toronto

We study the perfect information Nash equilibrium between a broker and her clients-an informed trader and an uniformed trader. In our model, the broker trades in the lit exchange where trades have instantaneous and transient price impact with exponential resilience, while both clients trade with the broker. The informed trader and the broker maximise expected wealth subject to inventory penalties, while the uninformed trader is not strategic and sends the broker random buy and sell orders. We characterise the Nash equilibrium of the trading strategies with the solution to a coupled system of forward-backward stochastic differential equations (FBSDEs). We solve this system explicitly and study the effect of information, profitability, and inventory control in the trading strategies of the broker and the informed trader.

Market Making with Exogenous Competition

R. Boyce, M. Herdegen, L. Sánchez-Betancourt
Oxford-Man Institute, University of Warwick, Imperial College London

We study liquidity provision in the presence of exogenous competition. We consider a 'reference market maker' who monitors her inventory and the aggregated inventory of the competing market makers. We assume that the competing market makers use a 'rule of thumb' to determine their posted depths, depending linearly on their inventory. By contrast, the reference market maker optimises over her posted depths, and we assume that her fill probability depends on the difference between her posted depths and the competition's depths in an exponential way. For a linear-quadratic goal functional, we show that this model admits an approximate closed-form solution. We illustrate the features of our model and compare against alternative ways of solving the problem either via an Euler scheme or state-of-the-art reinforcement learning techniques.

Unwinding Toxic Flow with Partial Information

A. Barzykin, R. Boyce, E. Neuman
Imperial College London

We consider a central trading desk which aggregates the inflow of clients' orders with unobserved toxicity, i.e. persistent adverse directionality. The desk chooses either to internalise the inflow or externalise it to the market in a cost effective manner. In this model, externalising the order flow creates both price impact costs and an additional market feedback reaction for the inflow of trades. The desk's objective is to maximise the daily trading PL subject to end of the day inventory penalization. We formulate this setting as a partially observable stochastic control problem and solve it in two steps. First, we derive the filtered dynamics of the inventory and toxicity, projected to the observed filtration, which turns the stochastic control problem into a fully observed problem. Then we use a variational approach in order to derive the unique optimal trading strategy. We illustrate our results for various scenarios in which the desk is facing momentum and mean-reverting toxicity. Our implementation shows that the P&L performance gap between the partially observable problem and the full information case are of order 0.01

Kullback-Leibler Barycentre of Stochastic Processes

S. Jaimungal, S. M. Pesenti
University of Toronto

We consider the problem where an agent aims to combine the views and insights of different experts' models. Specifically, each expert proposes a diffusion process over a finite time horizon. The agent then combines the experts' models by minimising the weighted Kullback-Leibler divergence to each of the experts' models. We show existence and uniqueness of the barycentre model and proof an explicit representation of the Radon-Nikodym derivative relative to the average drift model. We further allow the agent to include their own constraints, which results in an optimal model that can be seen as a distortion of the experts' barycentre model to incorporate the agent's constraints. Two deep learning algorithms are proposed to find the optimal drift of the combined model, allowing for efficient simulations. The first algorithm aims at learning the optimal drift by matching the change of measure, whereas the second algorithm leverages the notion of elicitability to directly estimate the value function. The paper concludes with a extended application to combine implied volatility smiles models that were estimated on different datasets.

Unified Approach for Hedging Impermanent Loss of Liquidity Provision

A. Lipton, V. Lucic, A. Sepp
Imperial College London

We develop static and dynamic approaches for hedging of the impermanent loss (IL) of liquidity provision (LP) staked at Decentralised Exchanges (DEXes) which employ Uniswap V2 and V3 protocols. We provide detailed definitions and formulas for computing the IL to unify different definitions occurring in the existing literature. We show that the IL can be seen a contingent claim with a non-linear payoff for a fixed maturity date. Thus, we introduce the contingent claim termed as IL protection claim which delivers the negative of IL payoff at the maturity date. We apply arbitrage-based methods for valuation and risk management of this claim. First, we develop the static model-independent replication method for the valuation of IL protection claim using traded European vanilla call and put options. We extend and generalize an existing method to show that the IL

protection claim can be hedged perfectly with options if there is a liquid options market. Second, we develop the dynamic model-based approach for the valuation and hedging of IL protection claims under a risk-neutral measure. We derive analytic valuation formulas using a wide class of price dynamics for which the characteristic function is available under the risk-neutral measure. As base cases, we derive analytic valuation formulas for IL protection claim under the Black-Scholes-Merton model and the log-normal stochastic volatility model. We finally discuss estimation of risk-reward of LP staking using our results.

Acknowledgement

The OMI Research Newsletter is curated by Álvaro Cartea, Director of the Oxford-Man Institute, and Fayçal Drissi, postdoctoral researcher at the institute. Paper suggestions are from fellow members at the institute: Álvaro Arroyo, Patrick Chang, Gerardo Durán-Martín, Marcello Monga, and Harrison Waldon.

References

Allmis, Patrick. 2024. "Cohesion, Ideology, and Tolerance." *arXiv preprint arXiv:2407.14045*.

Anthropelos, Michail, Runhuan Feng, and Seongyoon Kim. 2023. "On the Expansion of Risk Pooling."

Barzykin, Alexander, Robert Boyce, and Eyal Neuman. 2024. "Unwinding Toxic Flow with Partial Information." *arXiv preprint arXiv:2407.04510*.

Boyce, Robert, Martin Herdegen, and Leandro Sánchez-Betancourt. 2024. "Market Making with Exogenous Competition." *arXiv preprint arXiv:2407.17393*.

Cartea, Álvaro, Sebastian Jaimungal, and Leandro Sánchez-Betancourt. 2024. "Nash Equilibrium between Brokers and Traders." *arXiv preprint arXiv:2407.10561*.

Chopra, Hershdeep. 2024. "Certifying Lemons." *arXiv preprint arXiv:2407.19814*.

Jaimungal, Sebastian, and Silvana M Pesenti. 2024. "Kullback-Leibler Barycentre of Stochastic Processes." *arXiv preprint arXiv:2407.04860*.

Karl, Florian, Lukas Malte Kemeter, Gabriel Dax, and Paulina Sierak. 2024. "Position: Embracing Negative Results in Machine Learning." *arXiv preprint arXiv:2406.03980*.

Kim, Alex, Maximilian Muhn, and Valeri Nikolaev. 2024. "Financial statement analysis with large language models." *arXiv preprint arXiv:2407.17866*.

Lipton, Alexander, Vladimir Lucic, and Artur Sepp. 2024. "Unified approach for hedging impermanent loss of liquidity provision." *arXiv preprint arXiv:2407.05146*.

Moreno-Pino, Fernando, Álvaro Arroyo, Harrison Waldon, Xiaowen Dong, and Álvaro Cartea. 2024. "Rough Transformers: Lightweight Continuous-Time Sequence Modelling with Path Signatures." *arXiv preprint arXiv:2405.20799*.

Rathore, Pratik, Weimu Lei, Zachary Frangella, Lu Lu, and Madeleine Udell. 2024. "Challenges in training PINNs: A loss landscape perspective." *arXiv preprint arXiv:2402.01868*.

Rincón-Zapatero, Juan Pablo. 2024. "What does a dynamic oligopoly maximize? The continuous time Markov case." *arXiv preprint arXiv:2407.20810*.

Saeedi, Maryam, Yikang Shen, and Ali Shourideh. 2024. "Getting the Agent to Wait." *arXiv preprint arXiv:2407.19127*.

Yang, Parley R, and Alexander Y Shestopaloff. 2024. "Low Volatility Stock Portfolio Through High Dimensional Bayesian Cointegration." *arXiv preprint arXiv:2407.10175*.

Zhang, Yuyang, Shahriar Talebi, and Na Li. 2024. "Learning Low-dimensional Latent Dynamics from High-dimensional Observations: Non-asymptotics and Lower Bounds." *arXiv preprint arXiv:2405.06089*.