VOL. 2, NO. 9

OMI Research Newsletter – June 2023



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

Noteworthy contributions in this month's edition include works from the OMI, University of Oxford, King's College, Universidad Carlos III, Queen Mary University, and University College London. In this work, the authors use data from Euronext Amsterdam that includes algorithm and member identification to build statistical models to describe how market participants trade. Here, the authors generalise constant function markets (CFMs) and show that revenue from liquidity provision is suboptimal in these markets. They propose new designs for decentralised markets where liquidity providers set the price of liquidity as a function of their beliefs and risk preferences. In this work, the authors use a deep learning approach based on a convolutional Transformer encoder and a monotonic decoder to estimate the filltimes of limit orders posted in different levels of the LOB. Finally, the authors in this work prove the global convergence of the deep Galerkin method for solving PDEs.

Machine Learning & Statistics

Statistical Predictions of Trading Strategies in Electronic Markets

Á CARTEA, S. N. COHEN, R. GRAUMANS, S. LABYAD, L. SÁNCHEZ-BETANCOURT, AND L. VAN VELDHUIJZEN

OXFORD-MAN INSTITUTE, KING'S COLLEGE LONDON, ALAN TURING INSTITUTE, AUTORITEIT FINANCIËLE MARKTEN

We build statistical models to describe how market participants choose the direction, price, and volume of orders. Our dataset, which spans sixteen weeks for four shares traded in Euronext Amsterdam, contains all messages sent to the exchange and includes algorithm identification and member identification. We obtain reliable out-of-sample predictions and report the top features that predict direction, price, and volume of orders sent to the exchange. The coefficients from the fitted models are used to cluster trading behaviour and we find that algorithms registered as Liquidity Providers exhibit the widest range of trading behaviour among dealing capacities. In particular, for the most liquid share in our study, we identify three types of behaviour that we call (i) directional trading, (ii) opportunistic trading, and (iii) market making, and we find that around one third of Liquidity Providers behave as market markers.

Deep Attentive Survival Analysis in Limit Order Books: Estimating Fill Probabilities with Convolutional-Transformers

Á ARROYO, Á CARTEA, F. MORENO-PINO, S. ZOHREN

OXFORD-MAN INSTITUTE, UNIVERSIDAD CARLOS III DE MADRID

One of the key decisions in execution strategies is the choice between a passive (liquidity providing) or an aggressive (liquidity taking) order to execute a trade in a limit order book (LOB). Essential to this choice is the fill probability of a passive limit order placed in the LOB. This paper proposes a deep learning method to estimate the filltimes of limit orders posted in different levels of the LOB. We develop a novel model for survival analysis that maps time-varying features of the LOB to the distribution of filltimes of limit orders. Our method is based on a convolutional Transformer encoder and a monotonic neural network decoder. We use proper scoring rules to compare our method with other approaches in survival analysis, and perform an interpretability analysis to understand the informativeness of features used to compute fill probabilities. Our method significantly outperforms those typically used in survival analysis

literature. Furthermore, we carry out a detailed statistical analysis of the fill probability of orders placed in the order book, including within the bid-ask spread, for assets with different queue dynamics and trading activity.

Global Convergence of Deep Galerkin and PINNs Methods for Solving Partial Differential Equations

D. Jiang, J. Sirignano, and S. N. Cohen
University of Oxford

Numerically solving high-dimensional partial differential equations (PDEs) is a major challenge. Conventional methods, such as finite difference methods, are unable to solve high-dimensional PDEs due to the curse-of-dimensionality. A variety of deep learning methods have been recently developed to try and solve high-dimensional PDEs by approximating the solution using a neural network. In this paper, we prove global convergence for one of the commonly-used deep learning algorithms for solving PDEs, the Deep Galerkin Method (DGM). DGM trains a neural network approximator to solve the PDE using stochastic gradient descent. We prove that, as the number of hidden units in the single-layer network goes to infinity (i.e., in the "wide network limit"), the trained neural network converges to the solution of an infinite-dimensional linear ordinary differential equation (ODE). The PDE residual of the limiting approximator converges to zero as the training time $\to +\infty$. Under mild assumptions, this convergence also implies that the neural network approximator converges to the solution of the PDE. A closely related class of deep learning methods for PDEs is Physics Informed Neural Networks (PINNs). Using the same mathematical techniques, we can prove a similar global convergence result for the PINN neural network approximators. Both proofs require analyzing a kernel function in the limit ODE governing the evolution of the limit neural network approximator. A key technical challenge is that the kernel function, which is a composition of the PDE operator and the neural tangent kernel (NTK) operator, lacks a spectral gap, therefore requiring a careful analysis of its properties.

Robust Detection of Lead-Lag Relationships in Lagged Multi-Factor Models

Y. Zhang, M. Cucuringu, A. Y. Shestopaloff, S. Zohren
Oxford-Man Institute, Queen Mary University of London, Memorial University of Newfoundland, The Alan Turing Institute

In multivariate time series systems, key insights can be obtained by discovering lead-lag relationships inherent in the data, which refer to the dependence between two time series shifted in time relative to one another, and which can be leveraged for the purposes of control, forecasting or clustering. We develop a clustering-driven methodology for the robust detection of lead-lag relationships in lagged multi-factor models. Within our framework, the envisioned pipeline takes as input a set of time series, and creates an enlarged universe of extracted subsequence time series from each input time series, by using a sliding window approach. We then apply various clustering techniques (e.g, K-means++ and spectral clustering), employing a variety of pairwise similarity measures, including nonlinear ones. Once the clusters have been extracted, lead-lag estimates across clusters are aggregated to enhance the identification of the consistent relationships in the original universe. Since multivariate time series are ubiquitous in a wide range of domains, we demonstrate that our method is not only able to robustly detect lead-lag relationships in financial markets, but can also yield insightful results when applied to an environmental data set.

Deep Learning for Solving and Estimating Dynamic Macro-Finance Models

B. Fan, E. Qiao, A. Jiao, Z. Gu, W. Li, and L. Lu
Massachusetts Institute of Technology, University of Pennsylvania, Princeton, University of Southern California, University of Pennsylvania

We develop a methodology that utilizes deep learning to simultaneously solve and estimate canonical continuous-time general equilibrium models in financial economics. We illustrate our method in two examples: (1) industrial dynamics of firms and (2) macroeconomic models with financial frictions. Through these applications, we illustrate the advantages of our method: generality, simultaneous solution and estimation, leveraging the state-of-art machine-learning techniques, and handling large state space. The method is versatile and can be applied to a vast variety of problems.

Robust and Scalable Bayesian Online Changepoint Detection

M. Altamirano, F-X. Briol, and J. Knoblauch
University College London, The Alan Turing Institute

This paper proposes an online, provably robust, and scalable Bayesian approach for changepoint detection. The resulting algorithm has key advantages over previous work: it provides provable robustness by leveraging the generalised Bayesian perspective, and also addresses the scalability issues of previous attempts. Specifically, the proposed generalised Bayesian formalism leads to conjugate posteriors whose parameters are available in closed form by leveraging diffusion score matching. The resulting algorithm is exact, can be updated through

simple algebra, and is more than 10 times faster than its closest competitor.

When Do Neural Nets Outperform Boosted Trees on Tabular Data?

University of Maryland, New York University

Tabular data is one of the most commonly used types of data in machine learning. Despite recent advances in neural nets (NNs) for tabular data, there is still an active discussion on whether or not NNs generally outperform gradient-boosted decision trees (GBDTs) on tabular data, with several recent works arguing either that GBDTs consistently outperform NNs on tabular data, or vice versa. In this work, we take a step back and ask, 'does it matter?' We conduct the largest tabular data analysis to date, by comparing 19 algorithms across 176 datasets, and we find that the 'NN vs. GBDT' debate is overemphasized: for a surprisingly high number of datasets, either the performance difference between GBDTs and NNs is negligible, or light hyperparameter tuning on a GBDT is more important than selecting the best algorithm. Next, we analyze 965 metafeatures to determine what properties of a dataset make NNs or GBDTs better-suited to perform well. For example, we find that GBDTs are much better than NNs at handling skewed feature distributions, heavy-tailed feature distributions, and other forms of dataset irregularities. Our insights act as a guide for practitioners to decide whether or not they need to run a neural net to reach top performance on their dataset. Our codebase and all raw results are available at this https URL.

The GNAR-edge model: A network autoregressive model for networks with time-varying edge weights

A. Mantziou, M. Cucuringu, V. Meirinhos, and G. Reinert

Oxford-Man Institute, Office for National Statistics, The Alan Turing Institute, University of Oxford

In multivariate time series systems, key insights can be obtained by discovering lead-lag relationships inherent in the data, which refer to the dependence between two time series shifted in time relative to one another, and which can be leveraged for the purposes of control, forecasting or clustering. We develop a clustering-driven methodology for the robust detection of lead-lag relationships in lagged multi-factor models. Within our framework, the envisioned pipeline takes as input a set of time series, and creates an enlarged universe of extracted subsequence time series from each input time series, by using a sliding window approach. We then apply various clustering techniques (e.g, K-means++ and spectral clustering), employing a variety of pairwise similarity measures, including nonlinear ones. Once the clusters have been extracted, lead-lag estimates across clusters are aggregated to enhance the identification of the consistent relationships in the original universe. Since multivariate time series are ubiquitous in a wide range of domains, we demonstrate that our method is not only able to robustly detect lead-lag relationships in financial markets, but can also yield insightful results when applied to an environmental data set.

Seeing is Believing: Brain-Inspired Modular Training for Mechanistic Interpretability

Z. Liu, E. Gan, and M. Tegmark

MIT

We introduce Brain-Inspired Modular Training (BIMT), a method for making neural networks more modular and interpretable. Inspired by brains, BIMT embeds neurons in a geometric space and augments the loss function with a cost proportional to the length of each neuron connection. We demonstrate that BIMT discovers useful modular neural networks for many simple tasks, revealing compositional structures in symbolic formulas, interpretable decision boundaries and features for classification, and mathematical structure in algorithmic datasets. The ability to directly see modules with the naked eye can complement current mechanistic interpretability strategies such as probes, interventions or staring at all weights.

Explaining the Uncertain: Stochastic Shapley Values for Gaussian Process Models

S. L. Chau, K. Muandet, and D. Sejdinovic

University of Adelaide, CISPA

We present a novel approach for explaining Gaussian processes (GPs) that can utilize the full analytical covariance structure present in GPs. Our method is based on the popular solution concept of Shapley values extended to stochastic cooperative games, resulting in explanations that are random variables. The GP explanations generated using our approach satisfy similar favorable axioms to standard Shapley values and possess a tractable covariance function across features and data observations. This covariance allows for quantifying explanation uncertainties and studying the statistical dependencies between explanations. We further extend our framework to the problem of predictive explanation, and propose a Shapley prior over the explanation function to predict Shapley values for new data based on previously computed ones. Our extensive illustrations demonstrate the effectiveness of the proposed approach.

Nowcasting with signature methods

S. N. Cohen, S. Lui, W. Malpass, G. Mantoan, L. Nesheim, Á de Paula, A. Reeves, C. Scott, E. Small, and L. Yang

Oxford-Man Institute, Office for National Statistics, University College London, Economic Statistics Centre of Excellence, Bank of England

Key economic variables are often published with a significant delay of over a month. The nowcasting literature has arisen to provide fast, reliable estimates of delayed economic indicators and is closely related to filtering methods in signal processing. The path signature is a mathematical object which captures geometric properties of sequential data; it naturally handles missing data from mixed frequency and/or irregular sampling – issues often encountered when merging multiple data sources – by embedding the observed data in continuous time. Calculating path signatures and using them as features in models has achieved state-of-the-art results in fields such as finance, medicine, and cyber security. We look at the nowcasting problem by applying regression on signatures, a simple linear model on these nonlinear objects that we show subsumes the popular Kalman filter. We quantify the performance via a simulation exercise, and through application to nowcasting US GDP growth, where we see a lower error than a dynamic factor model based on the New York Fed staff nowcasting model. Finally we demonstrate the flexibility of this method by applying regression on signatures to nowcast weekly fuel prices using daily data. Regression on signatures is an easy-to-apply approach that allows great flexibility for data with complex sampling patterns.

Non-adversarial training of Neural SDEs with signature kernel scores

Z. Issa, B. Horvath, M. Lemercier, and C. Salvi

Oxford-Man Institute, King's College London, Imperial College London

Neural SDEs are continuous-time generative models for sequential data. State-of-the-art performance for irregular time series generation has been previously obtained by training these models adversarially as GANs. However, as typical for GAN architectures, training is notoriously unstable, often suffers from mode collapse, and requires specialised techniques such as weight clipping and gradient penalty to mitigate these issues. In this paper, we introduce a novel class of scoring rules on pathspace based on signature kernels and use them as objective for training Neural SDEs non-adversarially. By showing strict properness of such kernel scores and consistency of the corresponding estimators, we provide existence and uniqueness guarantees for the minimiser. With this formulation, evaluating the generator-discriminator pair amounts to solving a system of linear path-dependent PDEs which allows for memory-efficient adjoint-based backpropagation. Moreover, because the proposed kernel scores are well-defined for paths with values in infinite dimensional spaces of functions, our framework can be easily extended to generate spatiotemporal data. Our procedure permits conditioning on a rich variety of market conditions and significantly outperforms alternative ways of training Neural SDEs on a variety of tasks including the simulation of rough volatility models, the conditional probabilistic forecasts of real-world forex pairs where the conditioning variable is an observed past trajectory, and the mesh-free generation of limit order book dynamics.

Mathematical Finance

Automated Market Makers Designs Beyond Constant Functions

Á Cartea Fayal Drissi, L. Sánchez-Betancourt, David Šiška,f, Łukasz Szpruch

Oxford-Man Institute, University of Edinburgh, King's College London

Popular automated market makers (AMMs) use constant function markets (CFMs) to clear the demand and supply in the pool of liquidity. A key drawback in the implementation of CFMs is that liquidity providers (LPs) are currently providing liquidity at a loss, on average. In this paper, we propose new designs for AMMs where the price of liquidity is determined by LPs who post quotes around the marginal exchange rate of the pool. We develop two new models, the constant quote volume model (CQV) and the fractional quote volume model (FQV). In our models, the price of liquidity depends on the LP's (i) choice of impact functions that determine how liquidity taking orders impact the marginal rate, (ii) tolerance to inventory risk, and (iii) views on the demand for liquidity. We use stochastic control to obtain quotes that maximise the LP's expected profit, while controlling exposure to inventory risk. Our models admit closed-form solutions and are computationally efficient. We show that given a trading function of a CFM, there are impact functions and quotes in a CQV that replicate the marginal rate dynamics and the execution costs in the CFM; furthermore, we demonstrate that these quotes are suboptimal in the CQV pool. Finally, we use transaction data from Binance and Uniswap v3 to showcase optimal strategies for LPs and the profitability of liquidity provision in a CQV model.

Risk Budgeting Allocation for Dynamic Risk Measures

S. Jaimungal, S. M. Pesenti, Y. F. Saporito, and R. S. Targino

Oxford-Man Institute, University of Toronto, EMAp FGV

We define and develop an approach for risk budgeting allocation – a risk diversification portfolio strategy – where risk is measured using a dynamic time-consistent risk measure. For this, we introduce a notion of dynamic risk contributions that generalise the classical Euler contributions and which allow us to obtain dynamic risk contributions in a recursive manner. We prove that, for the class of dynamic coherent distortion risk measures, the risk allocation problem may be recast as a sequence of strictly convex optimisation problems. Moreover, we show that any self-financing dynamic risk budgeting strategy with initial wealth of 1 is a scaled version of the unique solution of the sequence of convex optimisation problems. Furthermore, we develop an actor-critic approach, leveraging the elicitability of dynamic risk measures, to solve for risk budgeting strategy using deep learning.

Calibration of Local Volatility Models with Stochastic Interest Rates using Optimal Transport

G. Loeper, J. Obloj and Benjamin Joseph

University of Oxford, BNP Paribas

We develop a non-parametric, optimal transport driven, calibration methodology for local volatility models with stochastic interest rate. The method finds a fully calibrated model which is the closest to a given reference model. We establish a general duality result which allows to solve the problem via optimising over solutions to a non-linear HJB equation. We then apply the method to a sequential calibration setup: we assume that an interest rate model is given and is calibrated to the observed term structure in the market. We then seek to calibrate a stock price local volatility model with volatility coefficient depending on time, the underlying and the short rate process, and driven by a Brownian motion which can be correlated with the randomness driving the rates process. The local volatility model is calibrated to a finite number of European options prices via a convex optimisation problem derived from the PDE formulation of semimartingale optimal transport. Our methodology is analogous to Guo, Loeper, and Wang, 2022 and Guo, Loeper, Obloj, et al., 2022a but features a novel element of solving for discounted densities, or sub-probability measures. We present numerical experiments and test the effectiveness of the proposed methodology.

Valuation of a Financial Claim Contingent on the Outcome of a Quantum Measurement

L. P. Hughston and L. Sánchez-Betancourt

University of London, King's College London, Oxford-Man Institute

We consider a rational agent who at time 0 enters into a financial contract for which the payout is determined by a quantum measurement at some time $T > 0$. The state of the quantum system is given by a known density matrix $\hat{p}$. How much will the agent be willing to pay at time 0 to enter into such a contract? In the case of a finite dimensional Hilbert space, each such claim is represented by an observable $\hat{X}_T$ where the eigenvalues of $\hat{X}_T$ determine the amount paid if the corresponding outcome is obtained in the measurement. We prove, under reasonable axioms, that there exists a pricing state $\hat{q}$ which is equivalent to the physical state $\hat{p}$ on null spaces such that the pricing function Π_{0T} takes the form $\Pi_{0T}(\hat{X}_T) = P_{0T}\,\mathrm{tr}(\hat{q}\hat{X}_T)$ for any claim $\hat{X}_T$, where P_{0T} is the one-period discount factor. By "equivalent" we mean that $\hat{p}$ and $\hat{q}$ share the same null space: thus, for any $|\xi\rangle \in \mathcal{H}$ one has $\langle\bar{\xi}|\hat{p}|\xi\rangle = 0$ if and only if $\langle\bar{\xi}|\hat{q}|\xi\rangle = 0$. We introduce a class of optimization problems and solve for the optimal contract payout structure for a claim based on a given measurement. Then we consider the implications of the Kochen-Specker theorem in such a setting and we look at the problem of forming portfolios of such contracts.

On Consistency of Signatures Using Lasso

X. Guo, R. Zhang, and C. Zhao

UC Berkeley, Peking University

Signature transforms are iterated path integrals of continuous and discrete-time time series data, and their universal nonlinearity linearizes the problem of feature selection. This paper revisits the consistency issue of Lasso regression for the signature transform, both theoretically and numerically. Our study shows that, for processes and time series that are closer to Brownian motion or random walk with weaker inter-dimensional correlations, the Lasso regression is more consistent for their signatures defined by Itô integrals; for mean reverting processes and time series, their signatures defined by Stratonovich integrals have more consistency in the Lasso regression. Our findings highlight the importance of choosing appropriate definitions of signatures and stochastic models in statistical inference and machine learning.

Exponential contractions and robustness for approximate Wonham filters

S. N. Cohen and E. Fausti
University of Oxford

We consider the problem of estimating the state of a continuous-time Markov chain from noisy observations. We show that the corresponding optimal filter is strictly contracting pathwise, when considered in the Hilbert projective space, and give explicit deterministic and pathwise rates of convergence. Using this, we provide alternative proofs of the robustness of optimal filters, improving on known error estimates, and derive rigorous and computable error bounds for approximate filters.

Economics

The Indoctrination Game

L. Ikan and D. Lagziel
Ben-Gurion University of the Negev

The indoctrination game is a complete-information contest over public opinion. The players exert costly effort to manifest their private opinions in public in order to control the discussion, so that the governing opinion is similar to theirs. Our analysis provides a theoretical foundation for the silent majority and vocal minority phenomena, i.e., we show that all moderate opinions remain mute in equilibrium while allowing extremists full control of the discussion. Moreover, we prove that elevated exposure to others' opinions increases the observed polarization among individuals. Using these results, we formulate a new social-learning framework, referred to as an indoctrination process.

Games Under Network Uncertainty

P. K. Chaudhuri, S. Sarangi, and H. Tzavellas
Virginia Tech

We consider an incomplete information network game in which agents' information is restricted only to the identity of their immediate neighbors. Agents form beliefs about the adjacency pattern of others and play a linear-quadratic effort game to maximize interim payoffs. We establish the existence and uniqueness of Bayesian-Nash equilibria in pure strategies. In this equilibrium agents use local information, i.e., knowledge of their direct connections to make inferences about the complementarity strength of their actions with those of other agents which is given by their updated beliefs regarding the number of walks they have in the network. Our model clearly demonstrates how asymmetric information based on network position and the identity of agents affect strategic behavior in such network games. We also characterize agent behavior in equilibria under different forms of ex-ante prior beliefs such as uniform priors over the set of all networks, Erdos-Renyi network generation, and homophilic linkage.

Mechanism Design without Rational Expectations

G. Rubbini
Brown University

Does dropping the rational expectations assumption mean the social planner can implement a larger class of social choice rules? This paper proposes a generalized model of implementation that does not assume rational expectations and characterizes the class of solution concepts requiring Bayesian Incentive Compatibility for full implementation. Surprisingly, full implementation of social choice functions turns out not to be significantly more permissive than with rational expectations. This implies some classical results, such as the impossibility of efficient bilateral trade (Myerson and Satterthwaite, 1983) hold for a broad range of non-equilibrium solution concepts, confirming their relevance even in boundedly rational setups.

Bubble Necessity Theorem

T. Hirano and A. A. Toda
University of London, University of California San Diego

Asset price bubbles are situations where asset prices exceed the fundamental values defined by the present value of dividends. This paper presents a conceptually new perspective on bubbles: the necessity of asset price bubbles. We establish the Bubble Necessity Theorem in a plausible general class of economic models: in economies with faster long run economic growth (G) than dividend growth (G_d) and long run bubbleless interest rate (R) below dividend growth, equilibria exist but none of them are fundamental or asymptotically bubbleless. The necessity condition $R < G_d < G$ naturally arises in models with uneven productivity growth and a sufficiently high savings motive.

Natural Language Processing

Distilling Step-by-Step! Outperforming Larger Language Models with Less Training Data and Smaller Model Sizes

C-Y. Hsieh, C-L. Li, C-K. Yeh, H. Nakhost, Y. Fujii, A. Ratner, R. Krishna, C-Y. Lee, and T. Pfister

University of Washington, Google Cloud AI Research, Google Research

Deploying large language models (LLMs) is challenging because they are memory inefficient and compute-intensive for practical applications. In reaction, researchers train smaller task-specific models by either finetuning with human labels or distilling using LLM-generated labels. However, finetuning and distillation require large amounts of training data to achieve comparable performance to LLMs. We introduce Distilling step-by-step, a new mechanism that (a) trains smaller models that outperform LLMs, and (b) achieves so by leveraging less training data needed by finetuning or distillation. Our method extracts LLM rationales as additional supervision for small models within a multi-task training framework. We present three findings across 4 NLP benchmarks: First, compared to both finetuning and distillation, our mechanism achieves better performance with much fewer labeled/unlabeled training examples. Second, compared to LLMs, we achieve better performance using substantially smaller model sizes. Third, we reduce both the model size and the amount of data required to outperform LLMs; our 770M T5 model outperforms the 540B PaLM model using only 80% of available data on a benchmark task.

RWKV: Reinventing RNNs for the Transformer Era

B. Peng, E. Alcaide, Q. Anthony, et al.

Ohio State University, RWKV Foundation, EleutherAI, University of Barcelona, Charm Therapeutics

Transformers have revolutionized almost all natural language processing (NLP) tasks but suffer from memory and computational complexity that scales quadratically with sequence length. In contrast, recurrent neural networks (RNNs) exhibit linear scaling in memory and computational requirements but struggle to match the same performance as Transformers due to limitations in parallelization and scalability. We propose a novel model architecture, Receptance Weighted Key Value (RWKV), that combines the efficient parallelizable training of Transformers with the efficient inference of RNNs. Our approach leverages a linear attention mechanism and allows us to formulate the model as either a Transformer or an RNN, which parallelizes computations during training and maintains constant computational and memory complexity during inference, leading to the first non-transformer architecture to be scaled to tens of billions of parameters. Our experiments reveal that RWKV performs on par with similarly sized Transformers, suggesting that future work can leverage this architecture to create more efficient models. This work presents a significant step towards reconciling the trade-offs between computational efficiency and model performance in sequence processing tasks.

Acknowledgement

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

References

Altamirano, Matias, François-Xavier Briol, and Jeremias Knoblauch. 2023. "Robust and Scalable Bayesian Online Changepoint Detection." *arXiv preprint arXiv:2302.04759*.

Arroyo, Álvaro, Álvaro Cartea, Fernando Moreno-Pino, and Stefan Zohren. 2023. "Deep Attentive Survival Analysis in Limit Order Books: Estimating Fill Probabilities with Convolutional-Transformers." *Available at SSRN 4432087*.

Cartea, Álvaro, Samuel N Cohen, Rob Graumans, Saad Labyad, Leandro Sánchez-Betancourt, and Leon van Veldhuijzen. 2023*a*. "Statistical Predictions of Trading Strategies in Electronic Markets." *Available at SSRN 4442770*.

Cartea, Álvaro, Fayçal Drissi, Leandro Sánchez-Betancourt, David Siska, and Lukasz Szpruch. 2023*b*. "Automated Market Makers Designs Beyond Constant Functions." *Available at SSRN 4459177*.

Chaudhuri, Promit K, Sudipta Sarangi, and Hector Tzavellas. 2023. "Games Under Network Uncertainty." *arXiv preprint arXiv:2305.03124*.

Chau, Siu Lun, Krikamol Muandet, and Dino Sejdinovic. 2023. "Explaining the Uncertain: Stochastic Shapley Values for Gaussian Process Models." *arXiv preprint arXiv:2305.15167*.

Cohen, Samuel N, Silvia Lui, Will Malpass, Giulia Mantoan, Lars Nesheim, Aureo de Paula, Andrew Reeves, Craig Scott, Emma Small, and Lingyi Yang. 2023. "Nowcasting with signature methods." *arXiv preprint arXiv:2305.10256*.

Fan, Benjamin, Edward Qiao, Anran Jiao, Zhouzhou Gu, Wenhao Li, and Lu Lu. 2023. "Deep Learning for Solving and Estimating Dynamic Macro-Finance Models." *arXiv preprint arXiv:2305.09783*.

Fausti, Eliana, and Samuel N Cohen. 2023. "Exponential contractions and robustness for approximate Wonham filters." *arXiv preprint arXiv:2305.02256*.

Guo, Xin, Ruixun Zhang, and Chaoyi Zhao. 2023. "On Consistency of Signatures Using Lasso." *arXiv preprint arXiv:2305.10413*.

Hirano, Tomohiro, and Alexis Akira Toda. 2023. "Bubble Necessity Theorem." *arXiv preprint arXiv:2305.08268*.

Hsieh, Cheng-Yu, Chun-Liang Li, Chih-Kuan Yeh, Hootan Nakhost, Yasuhisa Fujii, Alexander Ratner, Ranjay Krishna, Chen-Yu Lee, and Tomas Pfister. 2023. "Distilling step-by-step! outperforming larger language models with less training data and smaller model sizes." *arXiv preprint arXiv:2305.02301*.

Hughston, Lane P, and Leandro Sánchez-Betancourt. 2023. "Valuation of a Financial Claim Contingent on the Outcome of a Quantum Measurement." *arXiv preprint arXiv:2305.10239*.

Ikan, Lotem, and David Lagziel. 2023. "The Indoctrination Game." *arXiv preprint arXiv:2305.02604*.

Issa, Zacharia, Blanka Horvath, Maud Lemercier, and Cristopher Salvi. 2023. "Non-adversarial training of Neural SDEs with signature kernel scores." *arXiv preprint arXiv:2305.16274*.

Jaimungal, Sebastian, Silvana M Pesenti, Yuri F Saporito, and Rodrigo S Targino. 2023. "Risk Budgeting Allocation for Dynamic Risk Measures." *arXiv preprint arXiv:2305.11319*.

Jiang, Deqing, Justin Sirignano, and Samuel N Cohen. 2023. "Global Convergence of Deep Galerkin and PINNs Methods for Solving Partial Differential Equations." *arXiv preprint arXiv:2305.06000*.

Liu, Ziming, Eric Gan, and Max Tegmark. 2023. "Seeing is Believing: Brain-Inspired Modular Training for Mechanistic Interpretability." *arXiv preprint arXiv:2305.08746*.

Loeper, Gregoire, Jan Obloj, and Benjamin Joseph. 2023. "Calibration of Local Volatility Models with Stochastic Interest Rates using Optimal Transport." *arXiv preprint arXiv:2305.00200*.

Mantziou, Anastasia, Mihai Cucuringu, Victor Meirinhos, and Gesine Reinert. 2023. "The GNAR-edge model: A network autoregressive model for networks with time-varying edge weights." *arXiv preprint arXiv:2305.16097*.

McElfresh, Duncan, Sujay Khandagale, Jonathan Valverde, Ganesh Ramakrishnan, Micah Goldblum, Colin White, et al. 2023. "When Do Neural Nets Outperform Boosted Trees on Tabular Data?" *arXiv preprint arXiv:2305.02997*.

Peng, Bo, Eric Alcaide, Quentin Anthony, Alon Albalak, Samuel Arcadinho, Huanqi Cao, Xin Cheng, Michael Chung, Matteo Grella, Kranthi Kiran GV, et al. 2023. "RWKV: Reinventing RNNs for the Transformer Era." *arXiv preprint arXiv:2305.13048*.

Rubbini, Giacomo. 2023. "Mechanism Design without Rational Expectations." *arXiv preprint arXiv:2305.07472*.

Zhang, Yichi, Mihai Cucuringu, Alexander Y Shestopaloff, and Stefan Zohren. 2023. "Robust Detection of Lead-Lag Relationships in Lagged Multi-Factor Models." *arXiv preprint arXiv:2305.06704*.