Quantitative empirical verification of the Lillo-Mike-Farmer hypothesis for the persistent order flow in the Tokyo Stock Exchange market

Yuki Sato, **Kiyoshi Kanazawa**
Kyoto University, Kyoto, Japan

Statistical physics aims to study the macroscopic behaviour of various physical systems, and econophysics is its subfield to analyse the financial market microstructure by using recent high-frequency data in alignment with the statistical-physics program. In this talk, we focus on the persistence of the market order flow in financial markets from microscopic data analyses.

In broad financial markets, market orders exhibit the long-range correlation (LRC): once a buy (sell) market order is observed, a buy (sell) order is likely to appear even in future. In addition, this phenomenon is quantitatively characterised by the power-law decay of the sign autocorrelation function (ACF).

Then, a natural question will arise: what is the microscopic origin of this phenomenon? One of the promising microscopic hypotheses is the order-splitting hypothesis, stating that several large investors are splitting large metaorders into a long run of small child orders. Based on this hypothesis, Lillo, Mike, and Farmer (LMF) proposed a microscopic model in the spirit of statistical physics. This theory strongly predicts that the power-law exponent γ in the sign ACF is directly linked with the microscopic power-law exponent α in the run-length distribution for the order splitting, such that $\gamma = \alpha - 1$. However, this statistical-physics prediction has not been quantitatively verified in the absence of appropriate microscopic datasets for almost 18 years.

In this talk, we solve this long-standing problem by analysing a large microscopic dataset of the Tokyo Stock Exchange (TSE) markets based on our latest preprint (Y. Sato and K. Kanazawa, arXiv: 2301.13505). Our dataset includes the virtual server identifiers (IDs), a unit of trading accounts on the TSE market. By aggregating the virtual server IDs, we can allocate effective trader IDs to track traders' behaviour at the level of individual accounts. First, we study the run-length distributions of the order splitting to measure their microscopic power-law exponent α. We then measure the macroscopic power-law exponent γ in the order-sign ACFs. Finally, we show the scatterplot between α and γ, showing excellent and quantitative agreement with the LMF theory. Our work has provided the first quantitative support to the LMF model as the minimal microscopic model for the LRC in the order flow.