Direct observation of high‒frequency traders' strategies and theoretical foundation for financial Brownian motion

Article · March 2017

4 authors, including:

Hideki Takayasu
Sony Computer Science Laboratories, Inc.
286 PUBLICATIONS   **5,802** CITATIONS

SEE PROFILE

Misako Takayasu
Tokyo Institute of Technology
224 PUBLICATIONS   **3,345** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Appearance of Unstable Monopoly State Caused by Selective and Concentrative Mergers in Business Networks View project

Project   Econo-Socio-Physics View project

Direct observation of high-frequency traders' strategies and theoretical foundation for financial Brownian motion

Kiyoshi Kanazawa[1,2,*], Takumi Sueshige[2], Hideki Takayasu[1,3], and Misako Takayasu[1,2]

[1]*Institute of Innovative Research, Tokyo Institute of Technology,*
4259 Nagatsuta-cho, Yokohama 226-8502, Japan
[2]*Department of Mathematical and Computing Sciences,*
Graduate School of Information Science and Engineering,
Tokyo Institute of Technology, 4259 Nagatsuta-cho, Yokohama 226-8502, Japan
[3]*Sony Computer Science Laboratories, 3-14-13, Higashi-Gotanda, Shinagawa-ku, Tokyo, 141-0022, Japan*

Brownian motion has been a pillar of statistical physics for more than a century, and recent high-frequency trading data have shed new light on microstructure of Brownian motion in financial markets. Though evidences of trend-following behaviour of traders were indirectly shown in such trading data, the microscopic model has not been established so far by direct observation of trajectories for individual traders. In this paper, we present a minimal microscopic model for financial Brownian motion through an intensive analysis of trajectory data for all individuals in a foreign exchange market. This model includes a novel empirical law quantifying traders' trend-following behaviour that can create the inertial motion in market prices over short durations. We present a systematic solution paralleling molecular kinetic theory to reveal mesoscopic and macroscopic dynamics of our model. Our model exhibits quantitative agreements with empirical results strongly supporting our analysis.

I. INTRODUCTION

In physics, the study of colloidal Brownian motion has a long history beginning with Einstein's famous paper [1] in 1905, and the understanding of its mechanism has been systematically developed in molecular kinetic theory [2–4]. Recently, experimental developments have enabled researchers to observe a single trajectory itself of Brownian motion [5, 6], which engages physicists in quantitative modelling of sub-micrometer systems, such as molecular motors [7]. Likewise, recent breakthroughs for computational technologies have enabled physicists to study microstructure of financial Brownian motion in detail. They have applied their knowledge beyond material science and into studying in particular price movements in financial markets for about a quarter century. There consequently appeared new approaches as mentioned below, in contrast to conventional mathematical finance where a priori theoretical dynamics is assumed [8, 9].

There are three physical approaches to financial Brownian motions as shown in Figure 1. The microscopic approach focuses on the dynamics of traders in the financial markets (Figure 1a). Traders correspond to molecules in the modelling of materials, which enables both a numerical and a theoretical analysis of the macroscopic motion of prices [10–13]. Another approach is based on macroscopic empirical analyses of price time series (Figure 1c) and the direct empirical modelling of price dynamics focusing on fat-tailed distributions and long-time correlations in volatility [14–20]. The third approach focuses on mesoscopic dynamics concerning order-books (Figure 1b), which are accumulated buy–sell orders initiated by traders in the price axis where deals occur either at the best bid (buy) or ask (sell) price defining the market prices. Numerical simulations of markets become available by introducing models of order-book dynam-

ics [21–23]. Recently, the mesoscopic approach developed considerably with the analysis of high frequency financial market data in which the whole history of orders is tractable using a direct analogy between order-books and colloids [24]. From the order-book data [24], the importance of inertia was reported for market price, implying the existence of market trends. The Langevin equation was then found to hold most of time showing that the fluctuation–dissipation relation can be extended to order-book dynamics. However, microscopic mechanisms behind market trends have not been clarified so far because direct information is required on individual traders' strategies. In the present paper, we analyse more informative order-book data in which traders can be identified in an anonymised way for each order so that we can estimate each trader's dynamics directly from the data.

Here we present a minimal model validated by direct observation of individual traders' trajectories for financial Brownian motion. We first report a novel statistical law on the trend-following behaviour of foreign exchange (FX) traders by tracking the trajectories of all individuals. We next introduce a corresponding microscopic model incorporated with the empirical law. We reveal the mesoscopic and macroscopic behaviour of our model by developing a parallel theory to that for molecular kinetics: Boltzmann-like and Langevin-like equations are derived for the order-book and the price dynamics, respectively. A quantitative agreement with empirical findings is finally presented without further assumptions.

II. MICROSCOPIC DATA ANALYSIS

We analysed the high-frequency FX data between the US dollar (USD) and the Japanese Yen (JPY) from the 5th 16.00 to the 10th 20.00 GMT September 2016 on
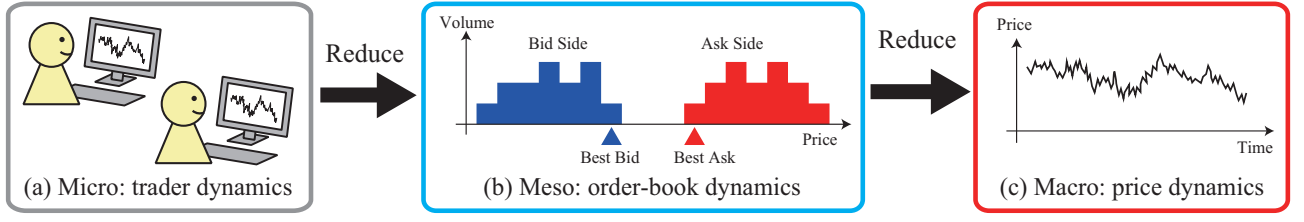
FIG. 1. Schematic of the hierarchy in the FX market from the viewpoint of (a) microscopic, (b) mesoscopic, and (c) macroscopic dynamics. The mesoscopic and macroscopic dynamics are obtained through a systematic reduction in the microscopic dynamics.
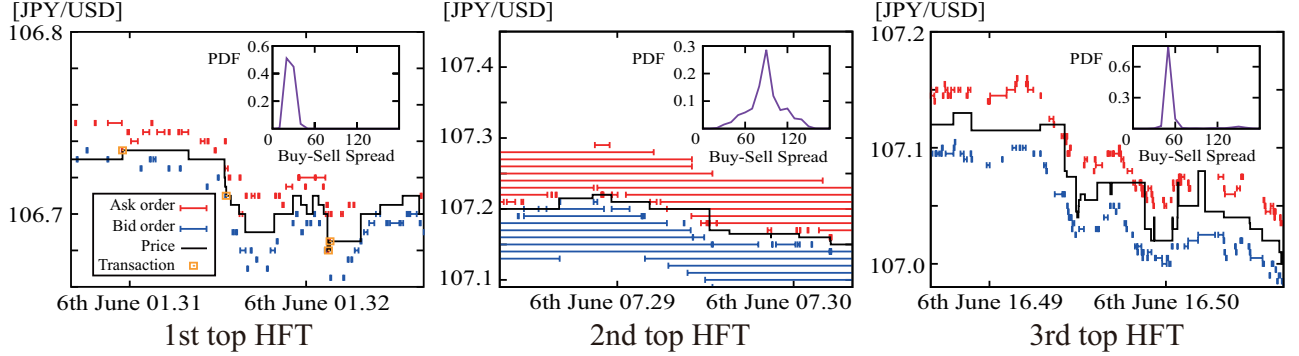

FIG. 2. Trajectories for the top 3 HFTs, where the lifetimes of orders are plotted for both bid and ask sides. Typical traders tend towards continuous two-sided quotes as market-makers, with the buy–sell spread fluctuating around a time-constant unique to the trader. The percentage of two-sided quotes among HFTs was 48.4% percents (see Appendix B).

Electronic Broking Services, one of the biggest FX platforms in the world. All trader activities were recorded for our dataset with anonymised trader IDs with one-millisecond time-precision. The minimum price-precision was 0.005 yen for the USD/JPY pair at that time, and the currency unit in this paper is 0.001 yen, called the tenth pip (tpip). The minimum volume unit for transaction was one million USD, and the total monetary flow was about 68 billion USD during this week. The market adopts the double-auction system, where traders quote bid or ask prices. In this paper, we particularly focused on the dynamics of high-frequency traders (HFTs), who frequently submit or cancel orders according to algorithms (see Appendix A for the definition). The presence of HFTs has rapidly grown recently [25] and 87.8% of the total orders were submitted by the HFTs in our dataset.

We first illustrate the trajectories of bid and ask prices quoted for the top 3 HFTs in Figure. 2a–c. We observed that with the two-sided quotes typical HFTs tend to play the role of liquidity providers (called market-makers [26, 27]). We also observed that buy–sell spreads (i.e., the difference between the bid and ask prices for a single market-maker) fluctuated around certain time-constants, showing a strong coupling between these prices. Indeed, the buy–sell spread distributions exhibit sharp peaks for individual HFTs as shown in the insets in Figure 2a–c (see also Appendix C).

We next report the empirical microscopic law for the trend-following strategy of individual traders. Let us denote the bid and ask prices of the top ith HFT by b_i and

a_i (see Appendix D for the definitions). We investigated the average movement of the mid-price $z_i \equiv (b_i + a_i)/2$ between transactions conditional on the previous market price movement Δp (Figure 3a). For the top 20 HFTs (Figure 3b and c), we find that the average movement is described by

$$\langle \Delta z_i \rangle_{\Delta p} \simeq c_i^* \tanh\left(\frac{\Delta p}{\Delta p_i^*}\right), \tag{1}$$

where the conditional average $\langle \ldots \rangle_{\Delta p}$ is taken when the previous price movement is Δp and $\Delta z_i \neq 0$ (see Appendix D for the detail). c_i^* and Δp_i^* are constants characterizing the price movement and the saturation threshold against the market trend. Here, typical values are given using $c_i^* \sim 4.2$ [tpip] and $\Delta p_i^* \sim 6$ [tpip]. The empirical law (1) implies that the reaction of traders is linear for small market trends but saturates for large market trends. Remarkably, a similar behaviour was reported from a full macroscopic analysis of market price data at one-month precision [28].

III. THEORETICAL MICROSCOPIC MODEL

Here we introduce a minimal microscopic model incorporating the empirical law (1). We make four assumptions: (i) The number of traders is sufficiently large. (ii) They always quote both bid and ask prices (for the ith trader, b_i and a_i) simultaneously with a unit vol-
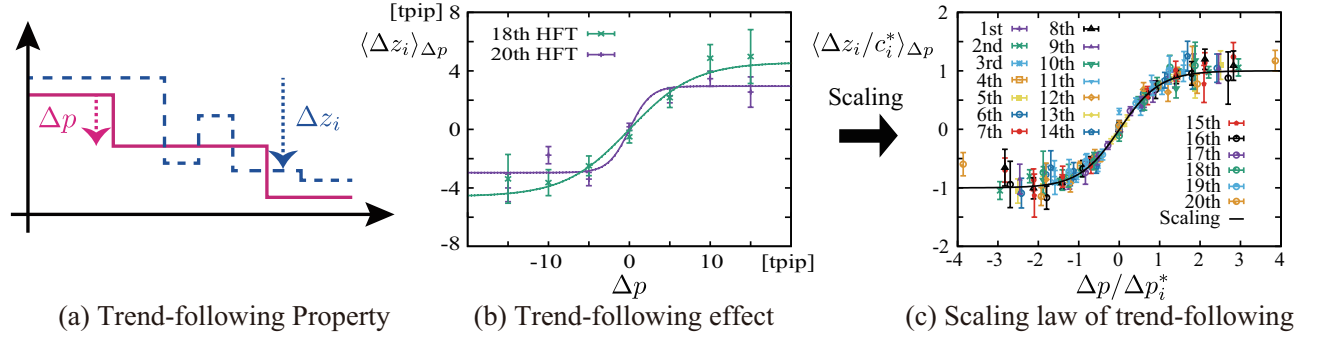
(a) Trend-following Property (b) Trend-following effect (c) Scaling law of trend-following

FIG. 3. (a) Quantification of trend-following for individual traders, where Δp and Δz_i are the movements of the market price and the mid-price of the ith trader, respectively. (b) Average Δz_i assuming previous price movements of Δp and an active trader with $\Delta z_i \neq 0$. The behaviour can be fitted by a hyperbolic function (1) with characteristic parameters Δp_i^* and c_i^*. (c) Curves for the top 20 HFTs showing trend-following scaled by the master curve (1).

(a) Trend-following random walk (b) Transaction (c) Requotation
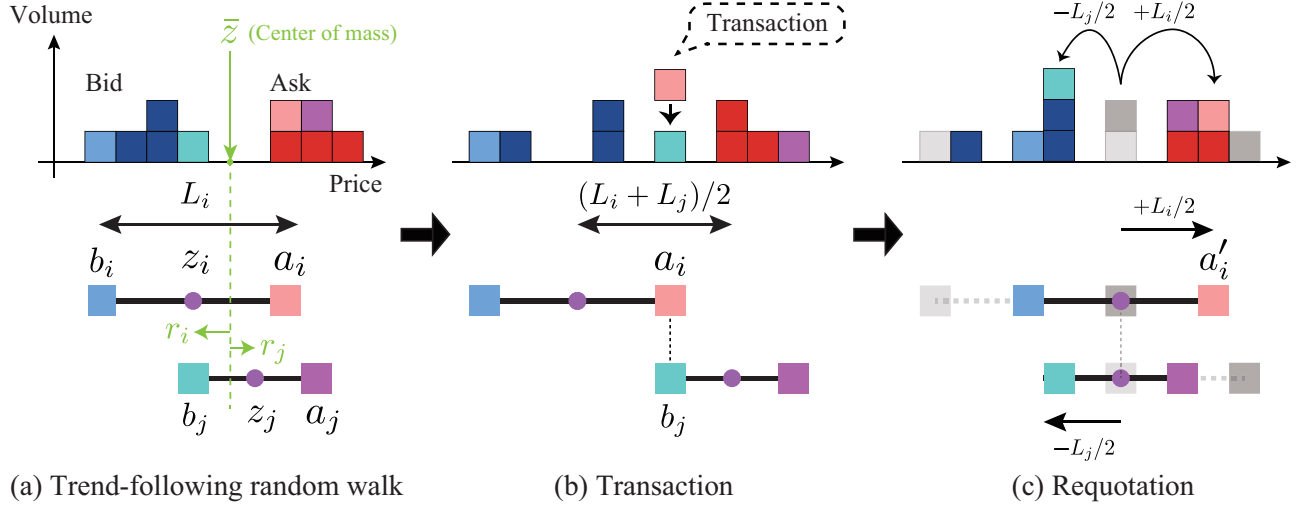
FIG. 4. Schematic of the dynamics of the model (2) and its correspondence to the order-book dynamics. (a) The mid-price of each trader moves according to trend-following and random movements. (b) When bid and ask prices match, a transaction takes place. (c) After the transaction, the pair of traders requote their bid and ask prices simultaneously.

ume as market-makers. (iii) Buy–sell spreads are time-constants unique to traders with distribution $\rho(L)$. The trader dynamics is then characterized by the mid-price $z_i \equiv (b_i + a_i)/2$. (iv) Trend-following random walks are assumed in the microscopic dynamics (see Figure 4a–c):

$$\frac{dz_i(t)}{dt} = c^* \tanh\left(\frac{\Delta p(t)}{\Delta p^*}\right) + \sigma^* \eta_i^{\mathrm{R}}(t) + \eta_i^{\mathrm{T}}(t) \qquad (2)$$

with a constant strength for trend-following c^*, white Gaussian noise $\sigma^* \eta_i^{\mathrm{R}}$ with constant variance σ^{*2}, and a requotation jump η_i^{T} after transactions (Figure 4c). η_i^{T} is defined by $\eta_i^{\mathrm{T}} \equiv \sum_{1 \leq j \leq N}^{j \neq i} \sum_{n=1} \Delta z_{ij} \delta(t - \tau_{n;ij})$ with jump size $\Delta z_{ij} = -L_i \mathrm{sgn}(z_i - z_j)/2$ and the nth transaction time $\tau_{n;ij}$ between traders i and j. The transaction condition at time t is given by $|z_i(t) - z_j(t)| = (L_i + L_j)/2$ (Figure 4b).

This model can be assessed using parallel tools employed in molecular kinetic theory. In this theory [2], the Boltzmann equation is first derived for the one-body

velocity distribution from the Hamiltonian dynamics by the method of Bogoliubov, Born, Green, Kirkwood, and Yvon assuming molecular chaos. The Langevin equation is derived in turn from the Boltzmann equation for massive Brownian particles [3, 4]. Here we have followed the same mathematical procedure to elucidate the dynamics behind order-book profiles and financial Brownian motion. For relative position r from the centre of mass $\bar{z} \equiv \sum_i z_i/N$, the Boltzmann-like equation is first derived for the one-body probability distribution density $\phi_L(r;t)$ conditional for a trader with a buy–sell spread L from the multi-agent dynamics (2):

$$\frac{\partial \phi_L}{\partial t} = \frac{\sigma^{*2}}{2} \frac{\partial \phi_L}{\partial r^2} + N \sum_{s=\pm 1} \int dL' \rho(L')[J_{LL'}^s(r+sL/2) - J_{LL'}^s] \qquad (3)$$

with $J_{LL'}^s(r) \equiv (\sigma^{*2}/2)|\partial_{rr'}|\{\phi_L(r)\phi_{L'}(r')\}|_{r-r'=s(L+L')/2}$ and $|\partial_{rr'}|f \equiv |\partial_r f| + |\partial_{r'} f|$. Here, $s = +1$ ($s = -1$) represents transactions as bid (ask) orders. The average order-

book profile is given by $f_A(r) = \int dL\rho(L)\phi_L(r - L/2)$ for the ask side. The integral term in equation (3) corresponds to the collision integral in the conventional Boltzmann equation. As the Langevin equation is derived from it [3], then similarly the Langevin-like equation is derived from the Boltzmann-like equation (3),

$$\Delta p(T + 1) = c^*\tau(T)\tanh(\Delta p(T)/\Delta p^*) + \zeta(T), \quad (4)$$

where $\Delta p(T)$ and $\tau(T)$ are the market price movement and transaction time interval at the Tth tick time. The tick time is an integer time incremented by every transaction and the mean time interval between transactions is 9.3 seconds in this week. The first and second terms on the right-hand side of Eq. (4) describe trend-following and random noise, respectively; the trend-following term corresponds to the momentum inertia in the conventional Langevin equation.

Equations (3) and (4) can be solved for $N \to \infty$ under an appropriate boundary condition (see Appendix E). We first set the buy–sell spread distribution as

$$\rho(L) = \frac{L^3}{6L^{*4}}e^{-L/L^*} \quad (5)$$

with decay length $L^* \simeq 23.5 \pm 0.5$ [tpip], empirically validated in our dataset (Figure 5a and Appendix C). The average order-book profile $f_A(r)$ is given for $r > 0$ by

$$f_A(r) = \frac{4e^{-\frac{3r}{2L^*}}}{3L^*}\left[\left(2 + \frac{r}{L^*}\right)\sinh\frac{r}{2L^*} - \frac{re^{-\frac{r}{2L^*}}}{2L^*}\right]. \quad (6)$$

The tail of the price movement is approximately given by

$$P(\geq |\Delta p|; \kappa) \sim e^{-|\Delta p|/\kappa} \quad (|\Delta p| \to \infty) \quad (7)$$

with decay length $\kappa \simeq 2\Delta z^*/3$, average movement from trend-following $\Delta z^* \equiv c^*\tau^*$, average transaction interval $\tau^* = 3L^{*2}/N\sigma^{*2}$, and complementary cumulative distribution $P(\geq |\Delta p|; \kappa)$. Further technical details are to be published in a forthcoming paper.

IV. MESOSCOPIC AND MACROSCOPIC DATA ANALYSIS

We next investigated the consistency between our microscopic model and our dataset. The empirical daily profile was first studied for the average order-book $f_A(r)$ in Figure 5b (see Appendix F for the detail). Surprisingly, we found an agreement with our theoretical lines (6) without fitting parameter that strongly supports the validity of our description.

We also empirically evaluated the two-hourly segmented cumulative distribution for the price movement in one-tick precision $P^{2h}(\geq |\Delta p|; \kappa)$ (Figure 5c), which obeys an exponential law that is consistent with our theoretical prediction (7). In our dataset, the decay length κ is approximately constant over a two-hour period but varies over time during the week. To remove this non-stationary feature, we introduced the two-hourly scaled cumulative distribution $\tilde{P}^{2h}(\geq |\Delta\tilde{p}|) \equiv P^{2h}(\geq \kappa|\Delta\tilde{p}|; \kappa)/Z$ with scaling parameters κ and Z (Figure 5d), thereby incorporating the two-hourly exponential-law for the whole week.

The price movements obey an exponential law for short periods but simultaneously obeys a power-law over long periods with exponent $\alpha = 3.6 \pm 0.13$ (Figure 5e). This apparent discrepancy is explained by the power-law nature of the decay length κ. Because κ approximately obeys a power-law cumulative distribution $Q(\geq \kappa) \sim \kappa^{-m}$ over the week with $m \simeq 3.5 \pm 0.13$ (Figure 5f), the one-week distribution $P^w(\geq |\Delta p|)$ obeys the power-law as a superposition of the two-hourly segmented exponential distribution,

$$P^w(\geq |\Delta p|) = \int_0^\infty d\kappa Q(\kappa)P^{2h}(\geq |\Delta p|; \kappa) \sim |\Delta p|^{-m} \quad (8)$$

with $Q(\kappa) \equiv -dQ(\geq \kappa)/d\kappa$. Remarkably, κ tends to be long when the market is inactive (Figure 5g). We therefore obtain a consistent result with the previously reported power-law [16–19] as a non-stationary property of κ.

We note that our model can exhibit super-diffusion under an appropriate parameter set as $\text{MSD}(T) \sim T^{2H}$ with $H \simeq 0.67 > 1/2$ for short periods (Figure 5h), which is consistent with previous reports [29]. Here the mean squared displacement is defined by $\text{MSD}(T) \equiv \langle(p(T) - \langle p(T)\rangle)^2\rangle$ as a function of tick time T with the ensemble average $\langle\ldots\rangle$. We also note that our model asymptotically shows ballistic behaviour $H = 1$ when trend-following is sufficiently large, which implies that it plays the role of a "momentum inertia" in financial markets. We further note that our model can show sub-diffusion ($H < 1/2$) when trend-following is sufficiently small.

V. CONCLUDING REMARKS

In this article, we have presented an intensive data analysis of anonymised traders in a foreign exchange market to directly observe strategies of HFTs. We first report a simple empirical law characterizing trend-following behaviour of individual traders against market trends. A trend-following random walk model is correspondingly introduced as a microscopic dynamics of the financial market. The mesoscopic and macroscopic behaviours of this model are systematically analysed in a parallel calculation to molecular kinetic theory. Our theoretical model reproduces the average order-book profile and the price movement distributions empirically. This work would be an important step toward unified description of financial markets from individual traders dynamics.

We discuss here a possible reason behind the success of our kinetic-like theory for our model. In material physics, mean-field approximations are invalid for low-dimensional systems because the low-dimensional geom-
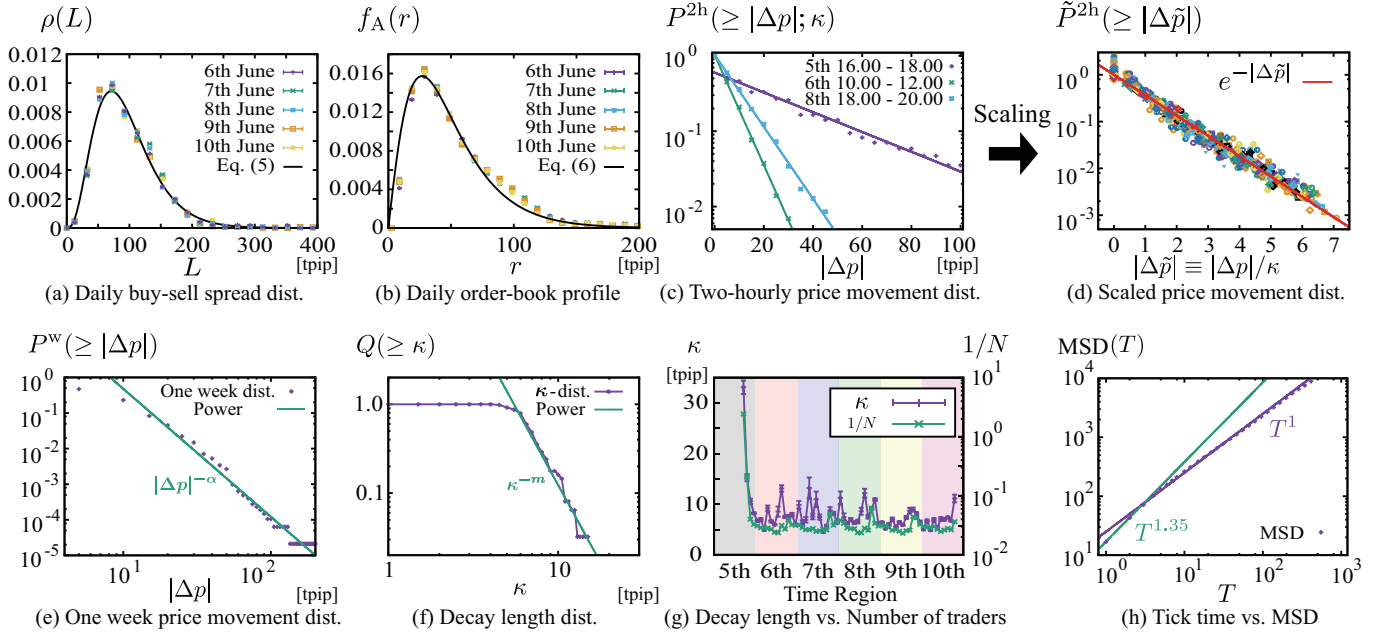
(a) Daily buy-sell spread dist.
(b) Daily order-book profile
(c) Two-hourly price movement dist.
(d) Scaled price movement dist.
(e) One week price movement dist.
(f) Decay length dist.
(g) Decay length vs. Number of traders
(h) Tick time vs. MSD

FIG. 5. (a) Daily distribution of buy–sell spreads for HFTs with the empirical master curve (5). (b) Daily average order-book profile for HFTs for the relative depth r from the market mid-price, which agrees with our theoretical line (6) without fitting parameters. (c) Two-hourly segmented cumulative distributions for the price movement in one-tick precision for the three typical time regions. The distributions are exponential, consistent with our theoretical prediction (7). (d) Two-hourly segmented cumulative distributions are scaled into the single exponential master curve every two-hours (62 time regions). (e) Price movement obeys a power-law over the whole week with a power-law exponent α. (f) Decay length κ obeys a power-law $Q(\geq \kappa) \sim \kappa^{-m}$ for $\kappa \to \infty$, theoretically implying a power-law behaviour in the price movement with exponent m over the week. This is consistent with the empirical power-law exponent $\alpha \simeq m$. (g) Correlation between the decay length κ and the number of HFTs N (See Appendix G). They are weakly correlated with Spearman's $\rho \simeq 0.3$. Remarkably, the decay length is longest in the inactive hours [27] (the 5th June 16.00–18.00 GMT, Sunday), which contributes strongly to the power-law tail. (h) Plot of the mean squared displacement $\mathrm{MSD}(T)$, which was obtained by a Monte Carlo simulation of our model with $L^* = 25$ [tpip], $\Delta p^* = 10$ [tpip], $\kappa = 5.25$ [tpip], and $N = 50$. This implies super-diffusion for a short time (i.e., $\mathrm{MSD}(T) \sim T^{2H}$ for small T with $H \simeq 0.67 > 1/2$) and normal diffusion over long times (i.e., $\mathrm{MSD}(T) \sim T^1$ for large T).

etry does not allow two-body correlations to disappear after collision. In contrast, in our model, traders are separated compulsorily after transactions, and there is little possibility for the same pair to enter successive transactions. The two-body correlation then quickly decays after collision assuming "molecular chaos" is valid. This scenario implies that the kinetic-like description may work well in various socio-economic systems, in addition to the previously studied examples, such as opinion formation and wealth distribution [30].

Our report dealt mainly with short-duration trends; their correlation with long-duration trends [28] is a topic to future studies. Of interest is the study of traders' be-

haviour in unstable markets triggered by external shocks, as those that occur in financial crises and flash crashes. The economics reason behind the hyperbolicity (1) in trend-following needs to be pursed further.

[1] Einstein, A. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Ann. Phys.* **322**, 549-560 (1905).

[2] Chapman, S. & Cowling, T. G. *The Mathematical Theory of Non-Uniform Gases* (Cambridge Univ. Press, Cambridge, 1970).

[3] Van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (Amsterdam, North-Holland, 2007).

[4] Van den Broeck, C., Kawai, R. & Meurs, P. Microscopic Analysis of a Thermal Brownian Motor. *Phys. Rev. Lett.* **93**, 090601-090604 (2004).

[5] Huang, R., Chavez, I., Taute, K. M., Lukić, B., Jeney, S., Raizen, M. G. & Florin, E.-L. Direct observation of the full transition from ballistic to diffusive Brownian motion in a liquid. *Nat. Phys.* **7**, 576-580 (2011).

[6] Li, T., Kheifets, S., Medellin, D. & Raizen, M. G. Measurement of the Instantaneous Velocity of a Brownian Particle. *Science* **328**, 1673-1675 (2010).

[7] Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. Jr Direct observation of the rotation of F1-ATPase. *Nature* **386**, 299-302 (1997).

[8] Black, F. & Scholes, M. S. The pricing of options and corporate liabilities. *J. Political Econ.* **81**, 637-654 (1973).

[9] Engle, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.* **20**, 339-350 (2002).

[10] Takayasu, H., Miura, H., Hirabayashi, T. & Hamada, K. Statistical properties of deterministic threshold elements – the case of market price. *Physica A* **184**, 127-134 (1992).

[11] Bak, P., Paczuski, M. & Shubik, M. Price variations in a stock market with many agents. *Physica A* **246**, 430-453 (1997).

[12] Lux, T. & Marchesi, M. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature* **397**, 498-500 (1999).

[13] Yamada, K., Takayasu, H., Ito, I. & Takayasu, M. Solvable stochastic dealer models for financial markets. *Phys. Rev. E* **79**, 051120-051131 (2009).

[14] Mantegna, R. N. & Stanley, H. E. Scaling behaviour in the dynamics of an economic index. *Nature* **376**, 46-49 (1995).

[15] Mantegna, R. N. & Stanley, H. E. *An Introduction to Econophysics* (Cambridge Univ. Press, 2000).

[16] Lux, T. The stable Paretian hypothesis and the frequency of large returns: an examination of major German stocks. *Appl. Financial Economics* **6** 463-475 (1996).

[17] Plerou, V., Gopikrishnan, P., Amaral, L. A., Meyer, M. & Stanley, H. E. Scaling of the distribution of price fluctuations of individual companies. *Phys. Rev. E* **60**, 6519-6528 (1999).

[18] Guillaume, D. M., Dacorogna, M. M., Davé, R. R., Müller, U. A., Olsen, R. B. & Pictet, O. V. From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance Stochast.* **1**, 95-129 (1997).

[19] Longin, F. M. The Asymptotic Distribution of Extreme Stock Market Returns. *J. Business* **69**, 383-408 (1996).

[20] Takayasu, M., Mizuno, T. & Takayasu, H. Potential force observed in market dynamics. *Physica A* **370**, 91-97 (2006).

[21] Maslov, S. Simple model of a limit order-driven market. *Physica A* **278**, 571-578 (2000).

[22] Smith, E., Farmer, J. D., Gillemot, L. & Krishnamurthy, S. Statistical theory of the continuous double auction. *Quantitative Finance* **3**, 481-514 (2003).

[23] Bouchaud, J.-P., Mézard, M. & Potters, M. Statistical properties of stock order books: empirical results and models. *Quantitative Finance* **2**, 251-256 (2002).

[24] Yura, Y., Takayasu, H., Sornette, D. & Takayasu, M. Financial Brownian Particle in the Layered Order-Book Fluid and Fluctuation-Dissipation Relations. *Phys. Rev. Lett.* **112**, 098703-098707 (2014).

[25] Hendershott, T., Jones, C. M. & Menkveld, A. J. Does Algorithmic Trading Improve Liquidity? *The Journal of Finance* **66**, 1-33 (2011).

[26] Menkveld, A. J. High frequency trading and the new market makers. *Journal of Financial Markets* **16**, 712-740 (2013).

[27] EBS Dealing Rules – Appendix – EBS Market (at the time of June 2016).

[28] Lempérière, Y., Deremble, C., Seager, P., Potters, M. & Bouchaud, J.-P. Two centuries of trend following. *Journal of Investment Strategies.* **3**, 41-61 (2014).

[29] Slanina, F. *Essentials of Econophysics Modelling.* (Oxford University, Oxford, 2014).

[30] Pareschi, L. & Toscani, G. *Interacting multiagent systems: kinetic equations and Monte Carlo methods.* (Oxford University Press, Oxford, 2013).

Appendix A: Definition of the high frequency traders

For this paper, we define a high frequent trader (HFT) as a trader who submits more than 500 times a day on average (i.e., more than 2500 times for the week). As several traders are unwilling to transact and often interrupt orders at the instant of submission (called flashing), we excluded traders with live orders of less than 0.5 percent of the transaction time. With this definition, the number of HFTs was 134 during this week, whereas the total number of traders was 1015. We note that the total number of traders who submitted limit orders was 922; the other 93 traders submitted only market orders.

Appendix B: Percentage of two-sided quotes

We calculated the percentage of two-sided quotes as follows: when a bid (ask) order is submitted by a trader, we check whether corresponding ask (bid) orders exist. We then count the number of two-sided quotes for all traders at every order submission and finally divide it by the total number of submissions.

Appendix C: Buy–sell spread

The difference in the median bid and ask prices was studied as a buy–sell spread for an HFT. Samples where only both bid and ask prices exist are taken in the one-second time-precision for Figure 2 and 5a. We plotted standard deviations of the averages as error bars for each point.

Appendix D: Trend-following effect

We remark on the precise definition of the bid (ask) price of individual HFTs for the analysis of trend-following. If a trader quotes both single-bid and single-ask orders at any time, the bid and ask prices are defined literally. In the presence of multiple bid or ask orders, we use the value of the median for the bid or ask orders as b_i or a_i. In the absence of any bid or ask orders, we use the previous bid or ask price as b_i or a_i for interpolation for Figure. 3b and c. We also remark that exceptional samples where the bid or ask price is far from the market price by 10 yen (0.0659% of the total) are excluded from the calculation of the conditional ensemble average $\langle\ldots\rangle_{\Delta p}$. We note that the standard deviations of the conditional averages are plotted for each point as error bars. Also, median values in the top 20 HFTs are given using $c_i^* \sim 4.2$ [tpip] and $\Delta p_i^* \sim 5.9$ [tpip].

Appendix E: Boundary conditions for the Boltzmann-like and Langevin-like equations

For the Boltzmann-like equation (3), we first introduce sufficiently large cutoffs at $r = \pm L_{\mathrm{cut}}/2$. The limit $N \to \infty$ is taken with reflecting boundaries assumed at $r = \pm L_{\mathrm{cut}}/2$. A large cutoff limit $L_{\mathrm{cut}} \to \infty$ is taken finally.

For the Langevin-like equation (4), we make the following two assumptions. (i) Trend-following has the same order as random noise: $c^* \tau_T \sim \zeta_T$. (ii) Saturation in

trend-following (1) is dominant: $\kappa \gg \Delta p^*$. We note that equation (1) can be approximated for large fluctuations as $\tanh(\Delta p/\Delta p^*) \simeq \mathrm{sgn}(\Delta p)$ under these conditions with the signature function $\mathrm{sgn}(x)$ defined by $\mathrm{sgn}(x) \equiv x/|x|$ for $x \neq 0$ and $\mathrm{sgn}(x) = 0$ for $x = 0$.

Appendix F: Average order-book profile

The daily average order-book profile is calculated for the HFTs. We took snapshots of the order-book every second and its ensemble average every day. We also plotted standard deviations of the averages as error bars for each point.

Appendix G: Number of Traders

We take snapshots of the order-book after every transaction and count the total number of different trader IDs for both bid and ask sides. The counting weight for an HFT quoting both sides is set to 1 and that for an HFT quoting one side is 1/2. We then plot the average of the number of trader IDs for both bid and ask sides every two hours in Figure 5g. The typical number of HFT IDs was about 35 in our dataset with this definition. The number of total volumes quoted by HFTs is typically about 80. Admittedly, there is room for debate on which number is appropriate for the calibration of the total number of traders in our model; it remains a topic for future study.