New Journal of Physics

The open access journal at the forefront of physics

PAPER

Time averaging, ageing and delay analysis of financial time series

Andrey G Cherstvy[1], Deepak Vinod[2], Erez Aghion[3], Aleksei V Chechkin[1,4] and Ralf Metzler[1,5]

[1] Institute for Physics & Astronomy, University of Potsdam, D-14476 Potsdam-Golm, Germany
[2] Dept of Inorganic & Physical Chemistry, Indian Institute of Science, 560012 Bangalore, India
[3] Institute of Nanotechnology and Advanced Materials, Bar Ilan University, Ramat-Gan 52900, Israel
[4] Institute for Theoretical Physics, Kharkov Institute of Physics and Technology, 61108 Kharkov, Ukraine
[5] Author to whom any correspondence should be addressed.

E-mail: a.cherstvy@gmail.com and rmetzler@uni-potsdam.de
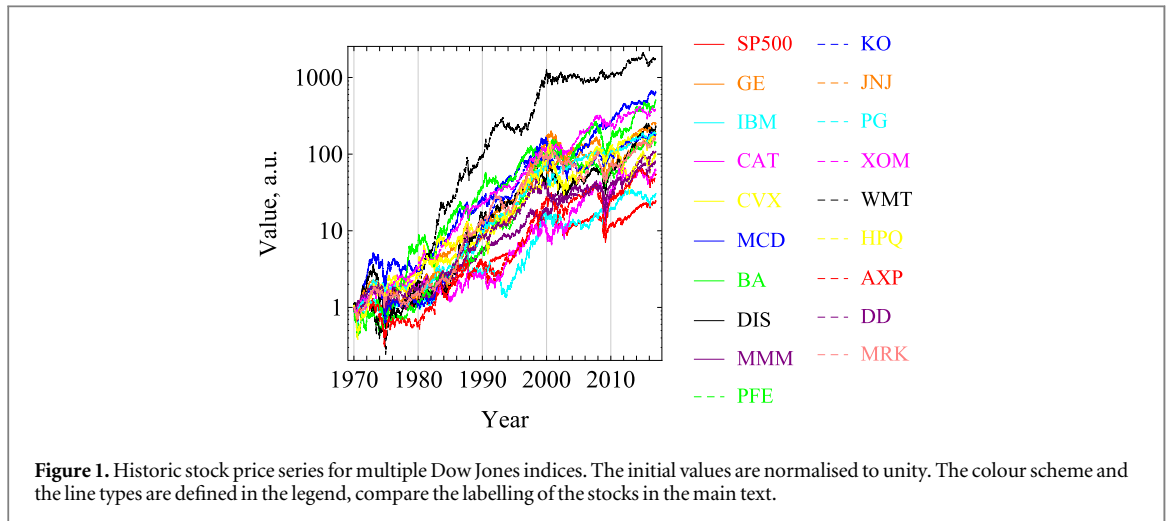
Abstract

We introduce three strategies for the analysis of financial time series based on time averaged observables. These comprise the time averaged mean squared displacement (MSD) as well as the ageing and delay time methods for varying fractions of the financial time series. We explore these concepts via statistical analysis of historic time series for several Dow Jones Industrial indices for the period from the 1960s to 2015. Remarkably, we discover a simple universal law for the delay time averaged MSD. The observed features of the financial time series dynamics agree well with our analytical results for the time averaged measurables for geometric Brownian motion, underlying the famed Black–Scholes–Merton model. The concepts we promote here are shown to be useful for financial data analysis and enable one to unveil new universal features of stock market dynamics.

1. Introduction

In 1900, Bachelier pioneered the concept that prices on financial markets are stochastic and may follow the laws of Brownian motion [1, 2]. Similar ideas for the mathematical description of option pricing were proposed by Bronzin in 1908 [3], a work which largely fell into oblivion until recently [4]. Later, in the 1960s, Mandelbrot [5, 6] and Fama [7] realised that the Gaussian distribution of price changes is often violated, and introduced Paretian and Lévy stable laws into financial mathematics [5, 8]. Another milestone in the stochastic modelling of financial markets is the famed Black–Scholes–Merton option pricing model [9–11], see also the classical textbooks on the mathematical analysis of markets [12–14].

The assessment of price variations for highly non-stationary [15–19] financial time series $X(t)$ (figure 1) is nowadays typically rationalised mathematically via solutions of stochastic differential equations with multiplicative noise [20–27]. This approach is based on the key concept of volatility: the magnitude of market fluctuations increases with the value $X(t)$. Numerous models for stock market price variation were developed to account for time-varying, clustered, and fluctuating volatility [12–14, 19, 26, 28–39]. In particular, the exponentially varying geometric Brownian motion (GBM) with its log-normal distribution, underlying the Black–Scholes–Merton model, is ubiquitously used in financial mathematics and econophysics [12, 14, 16, 40]. The GBM approach was generalised and extended for some alternative [24], specifically subdiffusive stochastic processes [41, 42], for power-law multiplicative noise [24–26, 43] (the square-root process [25, 44]), two-stock options [45], as well as fractional Brownian motion [46–48].

Statistical models of stock price variations often assume that the increments $\log[X(t)]$ are Gaussian-distributed [14]. Extensive statistical analysis of financial data revealed, however, that the distribution of returns $\log[X(t + dt)/X(t)]$ has a sharper maximum and fatter tails [8, 12, 19, 34]. To account for these features, random walk models based on, inter alia, truncated Lévy stable distributions [5, 8, 12, 36, 49] and jump-diffusion models [11, 13, 50] were proposed. Discrepancies between ensemble and time averaged measures were also discussed recently [18, 38, 51, 52]. Different market impacts onto price formation were considered as well [53]—in particular, one should mention here also market microstructure effects [54]. Despite decades of intense

Figure 1. Historic stock price series for multiple Dow Jones indices. The initial values are normalised to unity. The colour scheme and the line types are defined in the legend, compare the labelling of the stocks in the main text.

development and significant advances the quest for a universal mathematical model of stock price dynamics is still open.

Here we propose three concepts, complementary to the well established methods, for an advanced analysis of financial time series. Specifically, these are the time averaged mean squared displacement (MSD), perfectly suited for the analysis of a single time series, as well as the ageing and delay time methods. As we demonstrate from statistical analyses of real financial time series, these approaches are highly useful and reveal universal features of the market dynamics, which may be relevant for the further development of financial market models.

The central concept promoted here is based on time averaging. Most frequently, in theoretical approaches the MSD $\langle X^2(t) \rangle$, defined as the ensemble average of $X^2(t)$ over many realisations of the stochastic process $X(t)$, is used to quantify the time evolution of the process. However, when dealing with a single or few but long time series $X(t)$, the time averaged MSD [55]
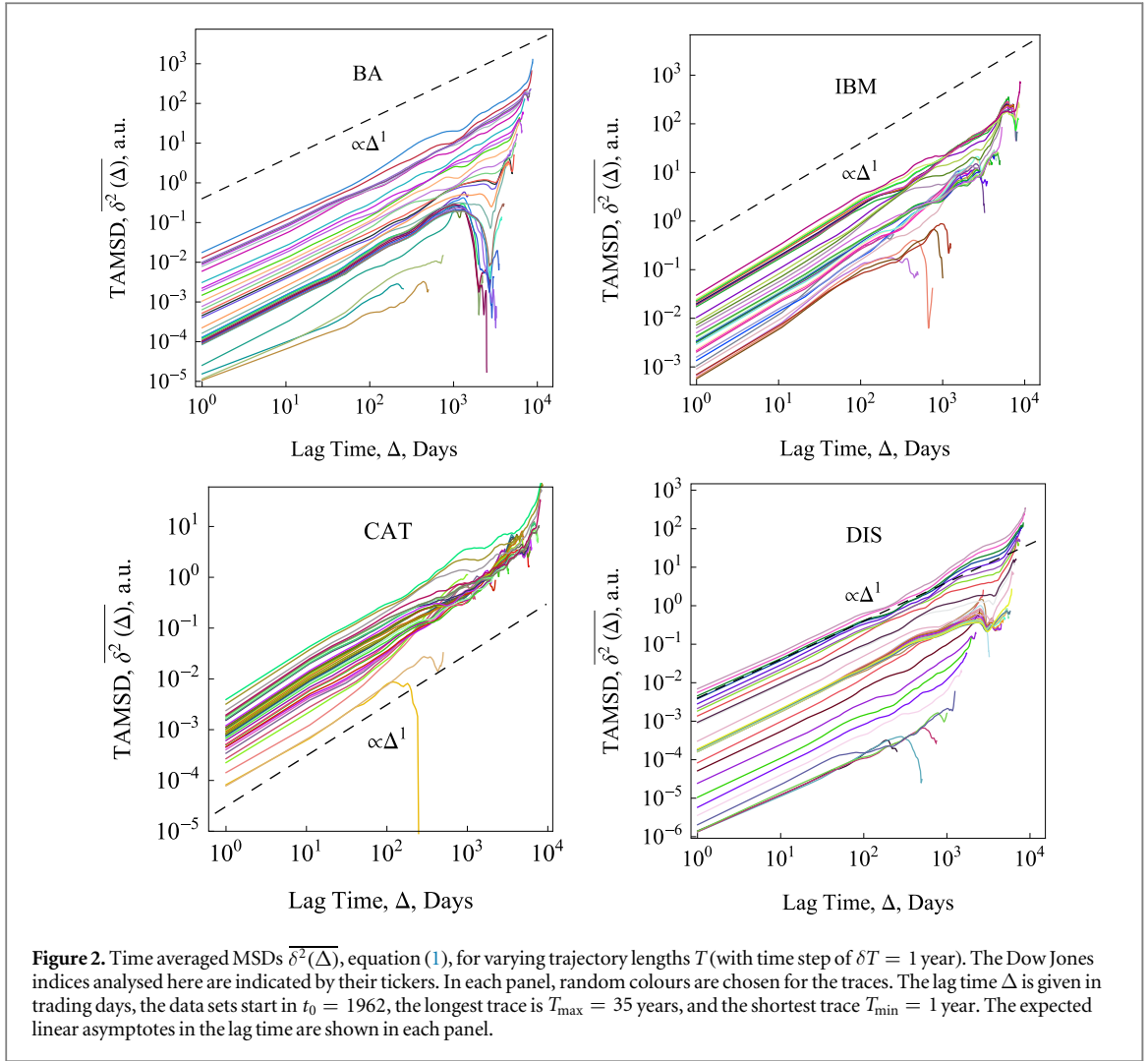
$$\overline{\delta^2(\Delta)} = \frac{1}{T - \Delta} \int_0^{T-\Delta} [X(t + \Delta) - X(t)]^2 \mathrm{d}t, \tag{1}$$

is better suitable and more relevant for the analysis of time series of both stationary and non-stationary stochastic processes [55, 56]. For a given lag time Δ this quantity defines a sliding average over the entire time series, of length T. Hereafter, the overline denotes time averages and the angular brackets stand for averages over an ensemble of realisations of a process $X(t)$. While the concept (1) is by now widely used for single-trajectory analysis in several areas of science, especially microscopic single particle tracking [55, 56], it is less common in mathematical finance [17, 18]. Our central focus here is to introduce this concept for the analysis of financial time series.

2. Analysis of financial time series

We here present the results of statistical analyses of financial time series based on the time averaged MSD (1) as well as the ageing and delay time methods defined below. The observed behaviour based on these analysis tools is demonstrated to agree well with analytical results for the famed GBM model, which is introduced in section 2.2. We study the daily values of multiple financial indices of different categories. All data were taken from and analysed via the `Wolfram Mathematica 9` platform.

The data are categorised and abbreviated as follows. In the first category we study the Dow Jones Industrial Average (`Mathematica` ticker symbol `^DJI`) group indices `SP500`: Standard and Poor's 500, `GE`: General Electric Comp, `IBM`: International Business Machines Corp, `CAT`: Caterpillar Inc, `CVX`: Chevron Corp, `MCD`: McDonald's Corp, `BA`: Boeing Comp, `DIS`: Walt Disney Comp, `MMM`: 3M Comp, `PFE`: Pfizer Inc, `KO`: Coca-Cola Comp, `JNJ`: Johnson and Johnson, `PG`: Procter and Gamble Comp, `XOM`: ExxonMobil Corp, `WMT`: Wal-Mart Stores, Inc, `AXP`: American Express Comp, `DD`: E. I. du Pont de Nemours and Comp, and `MRK`: Merck and Co Inc. In the next category we analyse some `DAX` indices at the Frankfurt Stock Exchange (Germany), with the tickers `BMW`: Bayerische Motoren Werke AG, `DAI`: DaimlerChrysler AG, `DTE`: Deutsche Telekom AG, `LHA`: Deutsche Lufthansa AG, `VOW3`: Volkswagen AG, `DPW`: Deutsche Post AG, `RWE`: RWE AG, `SAP`: SAP AG. Some high-tech tickers include `T`: AT&T Inc, `INTC`: Intel Corp, `HPQ`: Hewlett-Packard Comp, `MSFT`: Microsoft Corp. Finally, bank tickers comprise `JPM`: J. P. Morgan Chase and Co, `BAC`: Bank of America Corp, `NYSE:BCS`: Barclays PLC, `NYSE:SAN`: Santander-Chile Bank.

Figure 2. Time averaged MSDs $\overline{\delta^2(\Delta)}$, equation (1), for varying trajectory lengths T (with time step of $\delta T = 1$ year). The Dow Jones indices analysed here are indicated by their tickers. In each panel, random colours are chosen for the traces. The lag time Δ is given in trading days, the data sets start in $t_0 = 1962$, the longest trace is $T_{\max} = 35$ years, and the shortest trace $T_{\min} = 1$ year. The expected linear asymptotes in the lag time are shown in each panel.

The raw data of the price variations of several high-capitalisation Dow Jones indices starting at $t_0 = 1970$ are shown in figure 1. To improve the presentation we divided all prices by the corresponding initial value, such that all traces start at unit value. This is a legitimate procedure as for a GBM-like process the initial price X_0 enters solely as a multiplicative factor (see below). Despite the common initial price, we observe large price fluctuations for different indices at later times, especially in the last decades. Roughly, an exponential increase is observed for all stocks.

2.1. Time averaged MSD

For sufficiently long time series the time averaged MSD (1) typically produces relatively smooth results as a function of lag time. This feature is immediately obvious from the plots of $\overline{\delta^2(\Delta)}$ for data shown in figure 2. We emphasise that in the calculation of the time averaged MSD, for any given lag time Δ the entire time series is used and thus the initial values X_0 always contribute to the magnitude of $\overline{\delta^2(\Delta)}$. The lag time increment is 10 days in figure 2, and we used the 'raw' time series in this analysis (no normalisation to $X(t)$ values at the initial or final point is performed). Shorter lag times are computationally more expensive but would allow for a better resolution of the initial growth of $\overline{\delta^2(\Delta)}$. The spread of the computed magnitudes of $\overline{\delta^2(\Delta)}$ at short lag times increases with T for growing stock market prices, as long as no market crises occur, as seen in figure 2.

We observe a roughly linear growth $\overline{\delta^2(\Delta)} \propto \Delta$ at short lag times, in perfect accordance with the analytical result (6) for GBM derived below. This linear scaling stands in stark contrast to the exponential growth of the ensemble averaged MSD $\langle X^2(t) \rangle$ of GBM seen in equation (5) below. Such a fundamental discrepancy between the time and ensemble averaged MSD of the process $X(t)$ is well known in the theory of stochastic processes and often referred to as weak ergodicity breaking [55, 56]. For financial time series the fact that $\overline{\delta^2(\Delta)} \propto \Delta$ allows the direct analysis of $X(t)$ instead of its logarithm.

When the lag time approaches the length T of the time series, it is obvious from definition (1) that the statistic is worsening. This intrinsic property of time averages is reflected by the increasing fluctuations of $\overline{\delta^2(\Delta)}$ seen in figure 2. Note that in the limit $\Delta \to T$ the time-ensemble averaged MSD, averaged over N independent

trajectories of the same process, under identical starting conditions and parameters,

$$\left\langle \overline{\delta^2(\Delta)} \right\rangle = \frac{1}{N} \sum_{i=1}^{N} \overline{\delta_i^2(\Delta)}, \tag{2}$$

is equal to $\langle [X(t) - X_0]^2 \rangle$ [55]. Note that the averaging over different trajectories in quantity (2) is necessary for our analytical calculations. In the analysis of the financial time series shown in the figures, single trajectory averages $\overline{\delta^2(\Delta)}$ are considered throughout.

2.2. Geometric Brownian motion

Before continuing with our analysis of the actual stock market data, we briefly digress and provide a primer on GBM, a paradigm process employed in standard models for stock price dynamics. As we show this model reproduces the essential features observed in the market data.

GBM is defined in terms of the stochastic differential equation with multiplicative noise

$$dX(t) = \mu X(t) dt + \sigma X(t) dW(t). \tag{3}$$

Here $X(t) > 0$ is the price at time t, μ denotes a drift, and σ is the volatility (μ and σ are set constant below). The volatility is connected to the square root of the diffusivity [55]. The increments dW of the Wiener process $W(t) = \int_0^t \xi(t') dt'$ are defined as white Gaussian noise $\xi(t)$ with zero mean. The price evolution, starting with the initial value $X_0 > 0$ at $t = 0$, is obtained from equation (3) by use of Itô's lemma [12, 14] as

$$X(t) = X_0 e^{(\mu - \sigma^2/2)t + \sigma W(t)}. \tag{4}$$

This process satisfies the log-normal distribution, emerging also in models of task success and income distribution [57]. The MSD shows the exponential growth

$$\langle X(t)^2 \rangle = X_0^2 e^{(2\mu + \sigma^2)t}. \tag{5}$$

Due to this (much) faster growth compared to the linear increase of the MSD with t for Brownian motion, GBM is therefore a superdiffusive process [55].

To calculate the time averaged properties of interest, we resort to the average (2) over N independent realisations of $X(t)$. We derive this quantity using the one and two point distribution functions of the Wiener process. For short lag times, $\Delta \ll T$, and in the absence of drift we find

$$\left\langle \overline{\delta^2(\Delta)} \right\rangle \sim X_0^2 (e^{\sigma^2 T} - 1) \frac{\Delta}{T} \sim \langle X(T)^2 \rangle \frac{\Delta}{T}. \tag{6}$$

As mentioned already, the non-equivalence of the ensemble and time averaged MSD, $\langle X^2(\Delta) \rangle \neq \langle \overline{\delta^2(\Delta)} \rangle$ in the limit $\Delta/T \ll 1$ indicates weak ergodicity breaking, known for other non-stationary diffusion processes [55]. The time averaged MSD (6) scales linearly with the lag time Δ, in contrast to the exponential growth of $\langle X(t)^2 \rangle$, but grows exponentially with the trace length T, due to the highly non-stationary character of GBM.
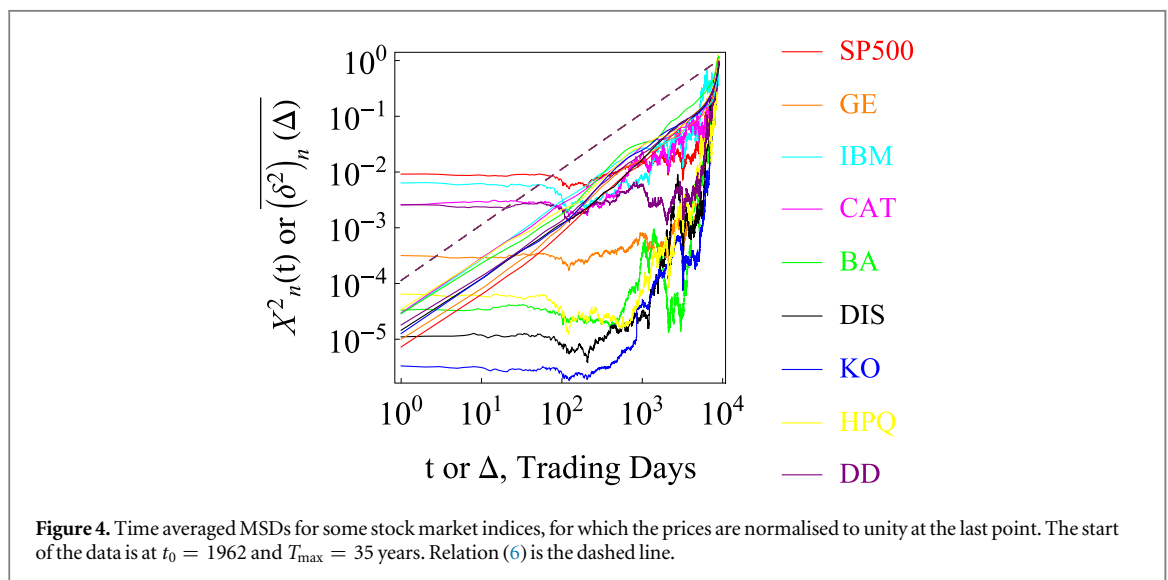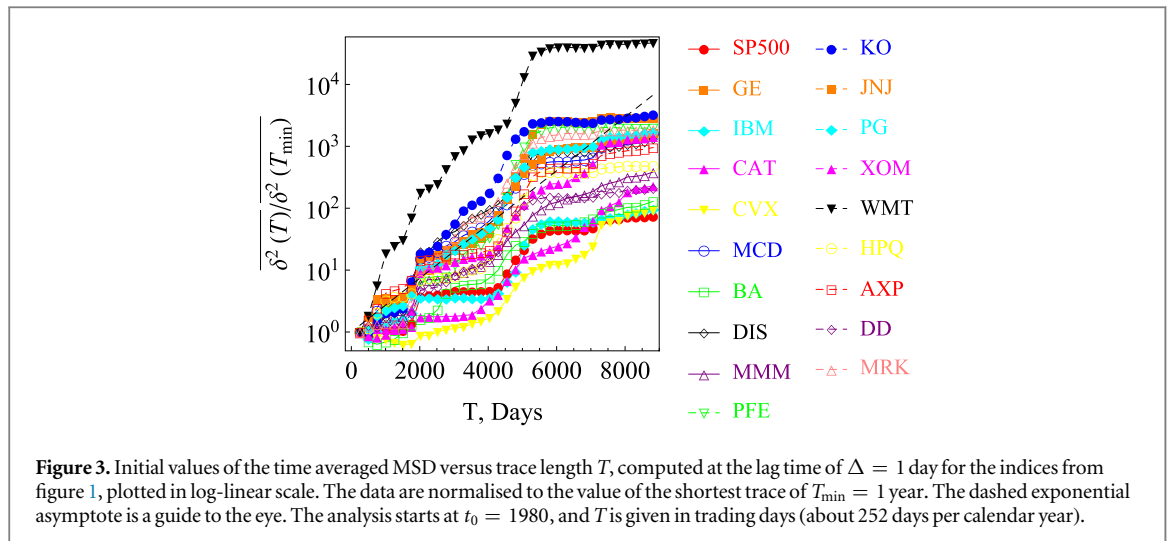
2.3. Time averaged MSD analysis of financial data, continued

We now return to the discussion of the time averaged MSD of the different stocks. The time averaged MSD of the analysed financial time series, similar to equation (6) for the GBM process, is expected to be a growing function of the trajectory length T. Indeed, this behaviour is demonstrated in figure 2, in which we plot the time averaged MSD $\overline{\delta^2(\Delta)}$ for varying trace length T for four different stock prices. This fact mirrors the accelerating nature of the stochastic process underlying the price variations.

What happens when the index value stagnates or drops substantially within a prolonged period of time, as can be seen in figure 1 following the 2008–2009 crisis? In this case the time averaged MSD $\overline{\delta^2(\Delta)}$ evaluated for partial time series (different T values) encompassing this time window tend to cluster together. Consequently, the value of $\overline{\delta^2(\Delta)}$ may saturate in this region, compare the original tick data in figure 1, the behaviour of $\overline{\delta^2(\Delta)}$ versus lag time Δ in figure 2, and the explicit dependence on T shown in figure 3.

Equation (6) predicts, to leading order, an exponential growth of the time-ensemble averaged MSD with the trace length T. As shown in figure 3, computed for $\Delta = 1$ day, the data indeed roughly follow an exponential increase with T. However, no universal dependence of the time averaged MSD as function of T could be found. Especially for those time windows encompassing prolonged drops or stalling of the index price, the ratio of the time averaged MSD for a trace length T to its value at $T_{\min}$ does not grow rapidly but rather saturates, as can also be seen in figure 3. Thus, checking the dependence of $\overline{\delta^2(\Delta)}$ on the trace length T may be used to unveil individual features of index price dynamics. In contrast to the strongly disparate and company-specific behaviour revealed in figure 3, universal features are found in the analysed index prices when employing the new concept of the delay time averaged MSD introduced below.

Figure 4 shows the data normalised to their end point. More specifically, this normalisation $\overline{\delta_n^2(\Delta)} \sim X_n^2(T)(\Delta/T)$ enhances the contribution of later parts of the time series, with typically larger prices.

Figure 3. Initial values of the time averaged MSD versus trace length *T*, computed at the lag time of $\Delta = 1$ day for the indices from figure 1, plotted in log-linear scale. The data are normalised to the value of the shortest trace of $T_{\min} = 1$ year. The dashed exponential asymptote is a guide to the eye. The analysis starts at $t_0 = 1980$, and *T* is given in trading days (about 252 days per calendar year).

Figure 4. Time averaged MSDs for some stock market indices, for which the prices are normalised to unity at the last point. The start of the data is at $t_0 = 1962$ and $T_{\max} = 35$ years. Relation (6) is the dashed line.

We find a considerable spread of individual $\overline{\delta_{i,n}^2}$ traces for different companies, because each index can have different parameters such as the volatility value or the attitude of a given company leadership to maximise the short-term profit versus ensuring a long-time sustainability. Finite lifetimes of companies and their varying age at the start of the time series are likewise important. These factors individualise stock price variations for each company and complicate the evaluation of ensemble averaged quantities. In view of this the universality of stock prices observed in section 2.5 is even more remarkable.
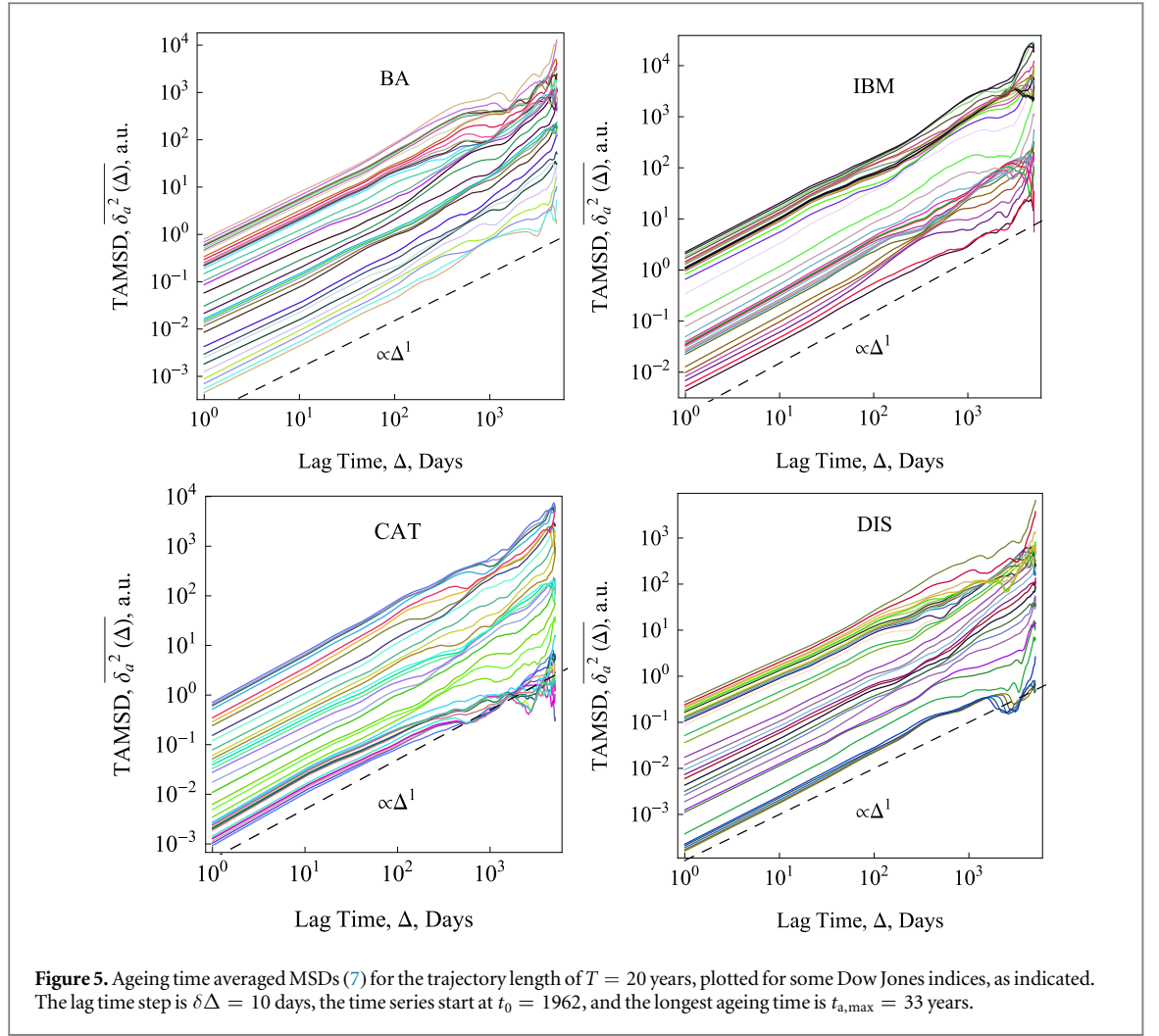
2.4. Ageing analysis

Physically, the above dependence (6) of the time averaged MSD on the trace length *T* reflects the phenomenon of ageing, a characteristic property of non-stationary stochastic processes [55, 58]. For superdiffusive processes with an MSD growing faster than the linear growth in time of Brownian motion, the time averaged MSD exhibits an increase with *T*. This reflects the self-reinforcing volatility of the process, as seen for the exponentially fast growth in equation (6). In contrast, in subdiffusion the effective diffusivity of the process is a decaying function of time [55]. For instance, in processes with scale free waiting times the typical sojourn periods of the motion become increasingly long, on average, effecting the decay of the time averaged MSD with *T* [58].

Another way to analyse ageing processes is the following. If the time series $X(t)$, starting at time $t = 0$, is evaluated only beginning with the so-called ageing time $t_a > 0$, the ageing time averaged MSD is defined as [58]

$$\overline{\delta_a^2(\Delta)} = \frac{1}{T - \Delta} \int_{t_a}^{t_a + T - \Delta} [X(t + \Delta) - X(t)]^2 \mathrm{d}t. \tag{7}$$

In this formal definition we shift the starting point for the analysis of the time series yet the length *T* of the analysed time interval remains fixed. Of course, larger values of t_a limit the remaining number of data points

Figure 5. Ageing time averaged MSDs (7) for the trajectory length of $T = 20$ years, plotted for some Dow Jones indices, as indicated. The lag time step is $\delta\Delta = 10$ days, the time series start at $t_0 = 1962$, and the longest ageing time is $t_{a,max} = 33$ years.

available for this analysis, however, the ageing time averaged MSD provides important insights into the underlying stochastic process [55]. We note that the term ageing does not imply any relaxation to an ergodic state in the limit of long ageing times, as it does for subdiffusive processes (the limit of strong ageing). In superdiffusive processes such as GBM and, as shown here, for highly non-stationary financial data the term ageing delineates the process time dependent increase of the spread of the random variable $X(t)$.
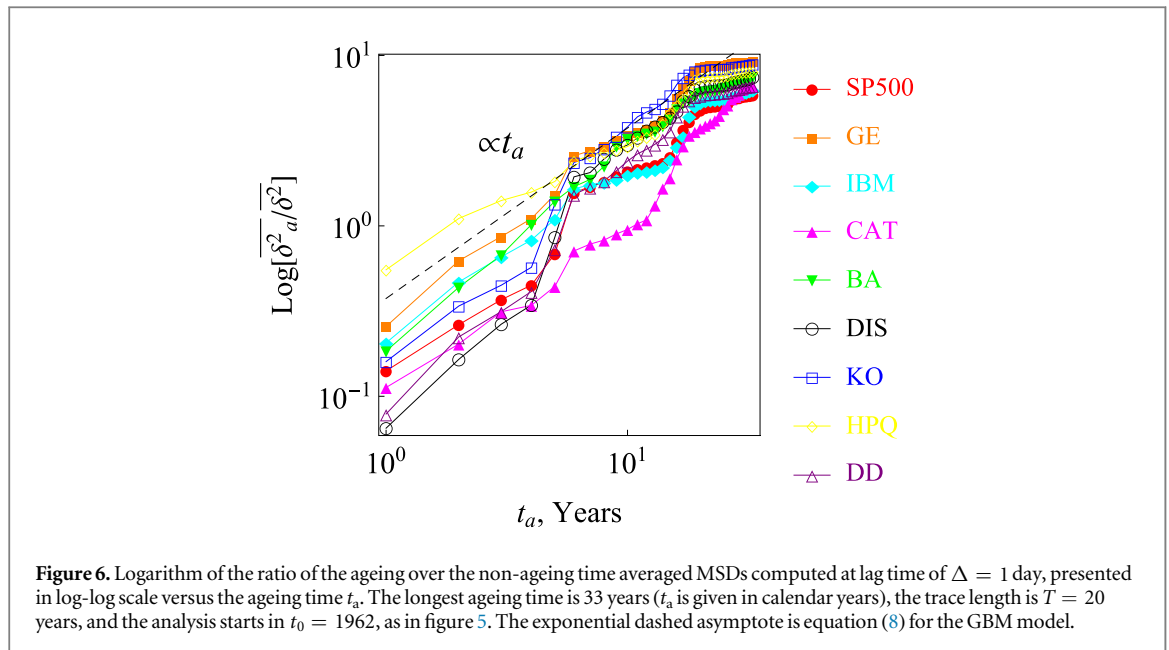
For GBM we find in the limit of short lag times, $\Delta \ll T$, and in absence of a drift, $\mu = 0$, that on average

$$\left\langle \overline{\delta_a^2(\Delta)} \right\rangle \sim \left\langle \overline{\delta^2(\Delta)} \right\rangle e^{\sigma^2 t_a}. \tag{8}$$

The exponential growth with t_a emerges as the process $X(t)$ already experienced an acceleration of the price dynamics up to the ageing time t_a, and thus the analysis starts with a higher volume. Note that in result (8) the argument of the exponential includes the volatility term σ^2 of the GBM process.

To see whether such ageing effects can indeed be observed in real financial data we study the behaviour of the ageing time averaged MSD $\overline{\delta_a^2(\Delta)}$ as function of the ageing time t_a. The raw data variation of $\overline{\delta_a^2(\Delta)}$ as function of the ageing time t_a for a number of indices is shown in figure 5. Indices growing rapidly and continuously in value reveal pronounced ageing effects, and $\overline{\delta_a^2(\Delta)}$ increases fast with t_a. Similar to the observations above, when we varied the trace length T, partial clustering of the traces for varying t_a is visible.

Figure 6 quantifies the behaviour of the logarithm of the ratio $\overline{\delta_a^2(\Delta)}/\overline{\delta^2(\Delta)}$ of the ageing time averaged MSD to the corresponding non-ageing value, plotted versus the ageing time t_a. Although, again, a roughly exponential increase is evident and consistent with the prediction (8) for GBM, no data collapse onto a universal curve is observed. Even at short ageing times we find a substantial spread of the $\log[\overline{\delta_a^2(\Delta)}/\overline{\delta^2(\Delta)}]$ curves versus t_a for different stock indices, as can be seen in figure 6. The non-universal behaviour is expected to be due to the fact that the volatility parameter varies between companies, as equation (8) predicts, and thus no data collapse in the growth of $\log[\overline{\delta_a^2(\Delta)}/\overline{\delta^2(\Delta)}]$ with the ageing time t_a occurs. No averaging over different companies is performed for the ageing time averaged MSD to compute the mean value $\left\langle \overline{\delta_a^2(\Delta)} \right\rangle$ in figure 6 because the

Figure 6. Logarithm of the ratio of the ageing over the non-ageing time averaged MSDs computed at lag time of $\Delta = 1$ day, presented in log-log scale versus the ageing time t_a. The longest ageing time is 33 years (t_a is given in calendar years), the trace length is $T = 20$ years, and the analysis starts in $t_0 = 1962$, as in figure 5. The exponential dashed asymptote is equation (8) for the GBM model.

corresponding ensemble is not formed from trajectories with identical parameters. For instance, the volatility values and the effect of the ageing time t_a for each stock market index can be markedly different.

2.5. Delay time analysis: revealing universal features

What if we allow the length of the time series to vary in the above ageing analysis? Namely, what would the expected behaviour be for the quantity

$$\overline{\delta_d^2(\Delta)} = \frac{1}{T - t_d - \Delta} \int_{t_d}^{T-\Delta} [X(t + \Delta) - X(t)]^2 \mathrm{d}t, \tag{9}$$

in which we call t_d the delay time, and the trace $X(t)$ is evaluated in the interval from $t = t_d$ until $t = T$. As long as the lag times Δ are short compared to the time span $T - t_d$ of the data, it can be shown that the mean delay time averaged MSD follows

$$\left\langle \overline{\delta_d^2(\Delta)} \right\rangle \sim \frac{\Delta}{T - t_d} X_0^2 (\mathrm{e}^{\sigma^2 T} - \mathrm{e}^{\sigma^2 t_d}). \tag{10}$$

For the ratio of the delay time averaged MSD (10) to the time averaged MSD (2) in the limit of short delay times $t_d / T \ll 1$ and long traces we get the simple, parameter-free result,
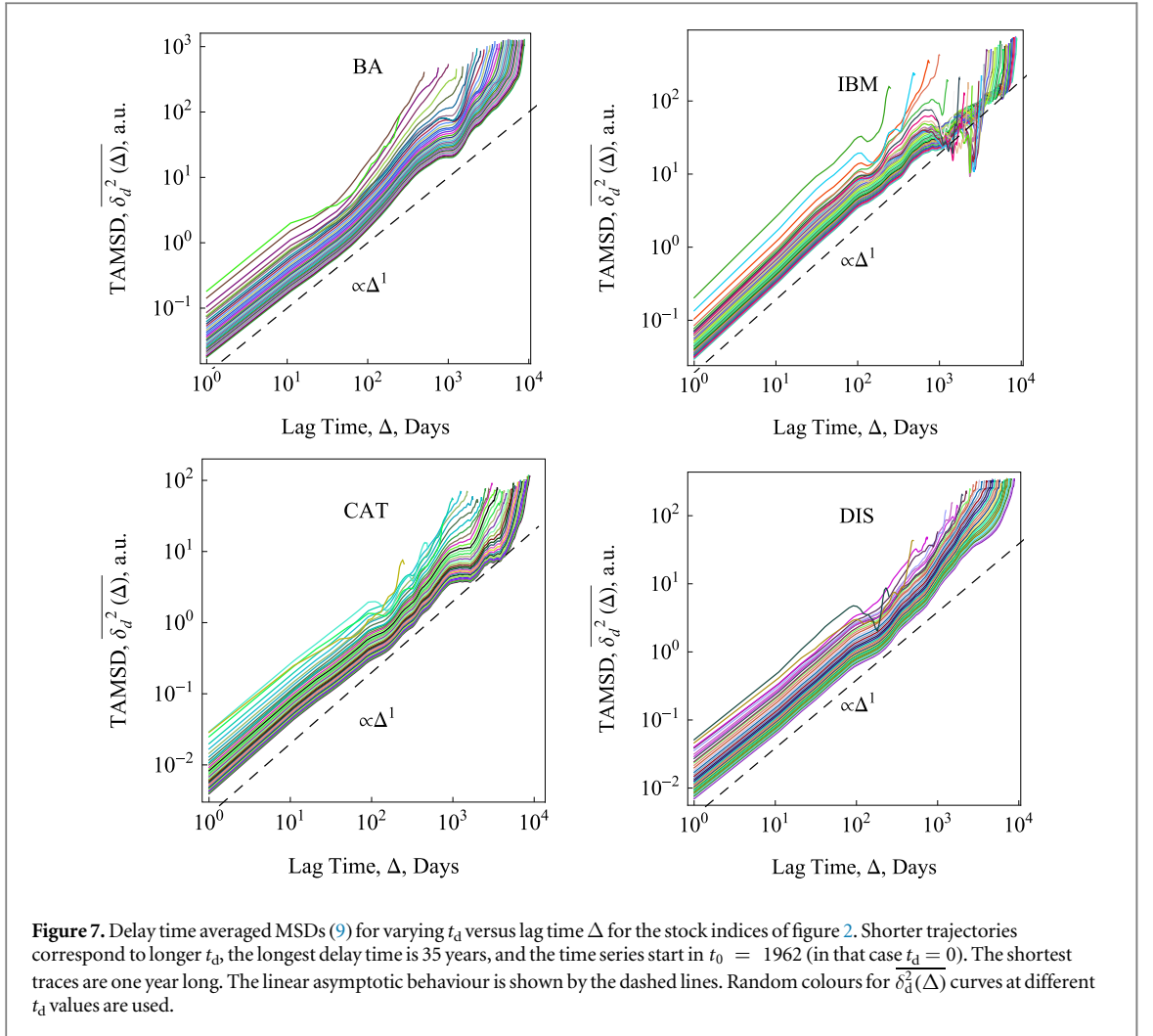
$$\log\left[\left\langle \overline{\delta_d^2(\Delta, t_d)} \right\rangle \Big/ \left\langle \overline{\delta^2(\Delta)} \right\rangle \right] \sim t_d / T. \tag{11}$$

The logarithm here cancels the leading exponential dependence of the delayed time averaged MSD on t_d in equation (10) and the final result (11) is independent on the index-specific volatility parameter σ. We emphasise that in this result it is crucial that the process $X(t)$ has non-stationary increments. The behaviour of, say, a standard Brownian process would follow a different scaling law with the delay time. Can this universal behaviour predicted in equation (11) indeed be seen in real financial time series, on the single trajectory level?

Pronounced up and down trends in the price evolution of an index give rise to strong variations of the $\overline{\delta_d^2(\Delta)}$ magnitudes, see figure 7. For longer t_d, the delay time averaged MSD includes a progressively shorter interval, causing a worsening statistic. Also, the magnitude of $\overline{\delta_d^2(\Delta)}$ increases because windows with higher prices and larger price variations are being processed in the averaging (9). For longer delay times t_d—when later parts of the time series contribute stronger to the time averaged MSD—the magnitude of $\overline{\delta_d^2(\Delta)}$ increases nearly linearly at short and moderate Δ, see figure 7. After a significant drop of stock index prices as a consequence of the 2008–2009 financial crisis, the variation of $\overline{\delta_d^2(\Delta)}$ with the delay time t_d exhibits large fluctuations for the lag times values encompassing this period of the time series. Due to this, the growing trend of $\overline{\delta_d^2(\Delta)}$ with the delay time may be reversed, see figure 7.

Figure 8 shows the logarithm of the ratio of the delay time averaged MSD to the standard time averaged MSD, evaluated at unit lag time $\Delta = 1$, as function of the delay time t_d. The universal behaviour (11) expected on average is followed very closely for each stock market time series. This universal behaviour—fulfilled for delay times t_d up to some 5–10 years—is the central result of this study. To our best knowledge, in terms of single time series this universal trend has not been reported before.

Figure 7. Delay time averaged MSDs (9) for varying t_d versus lag time Δ for the stock indices of figure 2. Shorter trajectories correspond to longer t_d, the longest delay time is 35 years, and the time series start in $t_0 = 1962$ (in that case $t_d = 0$). The shortest traces are one year long. The linear asymptotic behaviour is shown by the dashed lines. Random colours for $\overline{\delta_d^2(\Delta)}$ curves at different t_d values are used.

At longer delay times, we observe a distinct crossover to a somewhat steeper t_d-dependence in figure 8, consistent with a law of the form

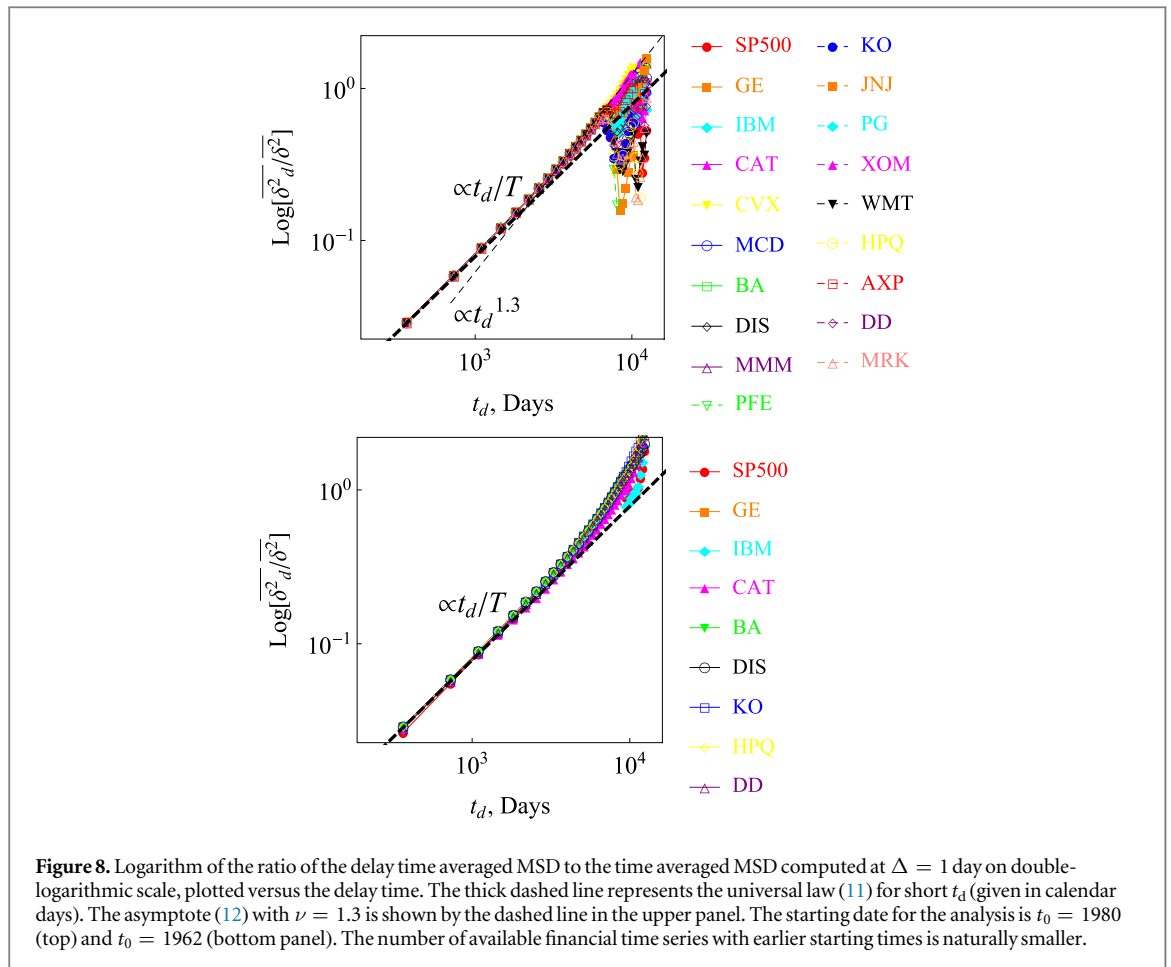$$\log\left[\overline{\delta_d^2(\Delta, t_d)}\middle/\overline{\delta^2(\Delta)}\right] \simeq (t_d/T)^\nu \tag{12}$$

with the scaling exponent $\nu > 1$. The observed behaviour features two relatively similar exponents: at short and intermediate delay times $\nu \approx 1$ while for long delay times $\nu \approx 1.3$. The initial linear scaling is universal, compare top and bottom panels of figure 8 with starting dates 1980 and 1962, respectively. As shown by the thin dashed line in the upper panel of figure 8, the second regime with $\nu > 1$ is more pronounced for the data starting $t_0 = 1980$, while it is less obvious for the $t_0 = 1962$ data. For long delay times we however consistently observe a faster than exponential growth of the ratio $\overline{\delta_d^2(\Delta, t_d)}/\overline{\delta^2(\Delta)}$. In particular, this disqualifies the predictions of GBM in this long delay time regime. Note that since the time averaged MSD initially grows approximately linearly with Δ, as one can see figures 2 and 7, the trends of figure 8 will also hold for other sufficiently short lag times.

Note that for some high tech companies and banks we did not observe a growth of $\overline{\delta_d^2(\Delta = 1, t_d)}/\overline{\delta^2(\Delta = 1)}$ with the delay time t_d (not shown). We believe that this is due to the very limited lengths of the time series available, starting in about 1987–1988. Therefore, effects of severe price drops during the 2008–2009 crisis dominate the magnitude of $\overline{\delta_d^2}$ over the entire time range we examined, rendering the result (11) for the standard GBM inapplicable. For a number of German DAX companies, with time series available from 2001, the prediction (12) does not hold either (not shown).

3. Conclusions

Time averaging of observables of a stochastic process $X(t)$ is a successful concept designed for the analysis of single or few, sufficiently long time series. It has been applied in various fields, in particular, in single particle tracking studies of microscopic objects [55, 56]. While for such microscopic particles, at least in principle, it is

Figure 8. Logarithm of the ratio of the delay time averaged MSD to the time averaged MSD computed at $\Delta = 1$ day on double-logarithmic scale, plotted versus the delay time. The thick dashed line represents the universal law (11) for short t_d (given in calendar days). The asymptote (12) with $\nu = 1.3$ is shown by the dashed line in the upper panel. The starting date for the analysis is $t_0 = 1980$ (top) and $t_0 = 1962$ (bottom panel). The number of available financial time series with earlier starting times is naturally smaller.

possible to record more than one trace under (almost) identical conditions, the situation is much more restricted for financial contexts. The market price evolution of a given company cannot be repeated several times under identical conditions. Thus, a statistical ensemble for averaging over a set of trajectories is inaccessible [17, 38]. Splitting up the time series into subparts is not an option due to the highly non-stationary character of the dynamics. The analysis in terms of time averaged observables is therefore the prime option.

Here we demonstrated that time averages indeed provide a useful toolbox for the analysis of financial data. Using the time averaged MSD, as well as the ageing and delay time methods, we showed that relevant features can be extracted from the analysis of financial time series. Good agreement of our data-driven observations with analytical predictions from the GBM model was observed, such as the linear lag time dependence of the time averaged MSD, contrasting the exponential growth for the ensemble averaged MSD. The ageing analysis combining the dependencies on the trace length T and the ageing time t_a unveil peculiar features in a given time series such as prolonged stalling or even a decrease of the stock value.

Remarkably, the delay time analysis introduced here uncovered a universal behaviour for the analysed stock prices. For short and intermediate delay times t_d, the logarithm of the ratio of the delay time versus the regular time averaged MSD is a linear function of t_d/T. At longer times, in our analysis beyond some 5–10 years, this logarithm approximately scales as a power-law $(t_d/T)^{\nu}$ with another scaling exponent, $\nu > 1$. This latter scaling behaviour is not captured with the standard GBM model, pointing at a need for improved theoretical approaches.

This study is to be viewed as a first step in applying time averaging, ageing, and delay time methods to the analysis of financial time series. Theoretically, modifications of the GBM model used here, to account for features such as the transition from the universal linear scaling to the scaling law $(t_d/T)^{\nu}$ for the delay time averaged MSD as well as the introduction of fluctuating or time-varying volatilities, are possible. The concept of 'diffusing diffusivities', which in some sense is similar to fluctuating volatilities, has been recently established in the physics literature [59–61]. How such concepts impact stochastic processes with multiplicative noise remains to be clarified, however, we expect a similarly rich behaviour with crossovers as observed for simple Brownian systems [59–61].

Deterministic time dependent volatilities may be adopted for the time averaging based description of unstable markets (at times of a financial crash), when the trading conditions change very rapidly [38]. Here, a

stochastic process with a power-law volatility may be proposed: GBM with a volatility increasing with time can account for a faster than exponential price growth $X(t)$ and explain a faster than linear trend (12) detected in the analysis of financial time series. Recently, ensemble averages of a similar modified GBM process with power-law and logarithmic volatilities were presented [53]. Also, models with value- and time-dependent diffusivity were empirically found to underlie the Euro–Dollar exchange rate dynamics [19, 62], compared to anomalous diffusion with space- and time-dependent diffusivity [55]. Combining such new theoretical approaches with time averaging may provide vital new impetus in the analysis of financial time series.

From a data analysis point of view, we were interested in the long-term trends for the time averaged MSD. Clearly, time series with one point a day hide possible intraday effects, such as intraday volatility patterns extracted from high-frequency data [18, 38, 63]. More observables from the time series should be taken into account, and the correlations between them remain to be rationalised. These include the auto-correlation function of price increments, the evaluation of volatility values [33] over different periods of time, the trading activity and volumes on the markets, a correction of the index price value due to inflation [34], crises, etc. These points, as well as the question how the dynamics observed herein is connected with heavy tailed spreads of financial volume [5, 8, 12, 36, 49] will be the focus of future work.

The area of mathematical finance is not the only domain where our time averaged MSD and ageing approaches may be useful. For instance, from a biological perspective the mathematical description of inherently highly stochastic disease outbreaks involves exponential processes similar to GBM. In epidemic spreading, an exponential increase in the number of diseased hosts is often postulated (up to the system size). The reader is referred to the optimal control models in epidemics spreading [64, 65], including density dependent growth and ageing. Finally, mathematical models of tumour spreading and the growth of bacterial colonies and cells [66] also employ exponential processes, providing additional ground for the application of the concepts outlined here.

Acknowledgments

ORCID

Ralf Metzler ⓘ https://orcid.org/0000-0002-6013-7020

References

[1] Bachelier L 1900 Théorie de la Spéculation *Ann. Ecole Norm. Super.* **17** 21
[2] Mandelbrot B B 1983 *The Fractal Geometry of Nature* (New York: Freeman)
[3] Bronzin V 1908 *Theorie der Prämiengeschäfte* (Leipzig und Wien: Verlag Franz Deticke)
[4] Hafner W and Zimmermann H 2009 *Vinzenz Bronzin's Option Pricing Models: Exposition and Appraisal* (Berlin: Springer)
[5] Mandelbrot B B 1963 The variation of certain speculative Prices *J. Bus.* **36** 394
[6] Mandelbrot B B 1963 New methods in statistical economics *J. Polit. Econ.* **71** 421
[7] Fama E F 1965 The behavior of stock-market prices *J. Bus.* **38** 34
[8] Mantegna R N and Stanley H E 1995 Scaling behaviour in the dynamics of an economic index *Nature* **376** 46
[9] Black F and Scholes M 1973 The pricing of options and corporate liabilities *J. Polit. Econ.* **81** 637
[10] Merton R C 1971 Optimum consumption and portfolio rules in a continuous-time model *J. Econ. Theory* **3** 373
[11] Merton R C 1976 Option pricing when underlying stock returns are discontinuous *J. Financ. Econ.* **3** 125
[12] Mantegna R N and Stanley H E 2000 *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge: Cambridge University Press)
[13] Bouchaud J-P and Potters M 2000 *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge: Cambridge University Press)
[14] Fouque J-P, Papanicolaou G and Sircar K R 2000 *Derivative in Financial Markets with Stochastic Volatility* (Cambridge: Cambridge University Press)
[15] McCauley J L 2009 *Dynamics of Markets: The New Financial Economics* (Cambridge: Cambridge University Press)
[16] McCauley J L 2013 *Stochastic Calculus and Differential Equations for Physics and Finance* (Cambridge: Cambridge University Press)
[17] McCauley J L 2008 Time versus ensemble averages for nonstationary time series *Physica* A **387** 5518
[18] Seemann L, Hua J-C, McCauley J L and Gunaratne G H 2012 Ensemble versus time averages in financial time series analysis *Physica* A **391** 6024
[19] Bassler K E, McCauley J L and Gunaratne G H 2007 Nonstationary increments, scaling distributions, and variable diffusion processes in financial markets *Proc. Natl Acad. Sci. USA* **104** 17287
[20] Feller W 1951 Two singular diffusion problems *Ann. Math.* **54** 173
[21] Samuelson P A 1965 Proof that properly anticipated prices fluctuate randomly *Ind. Manage. Rev.* **6** 41
[22] Duffie D 2001 *Dynamic Asset Pricing Theory* 3rd edn (Princeton, NJ: Princeton University Press)
[23] Øksendal B 2005 *Stochastic Differential equations* 6th edn (Berlin: Springer)
[24] Cox J C and Ross S A 1976 The valuation of options for alternative stochastic processes *J. Financ. Econ.* **3** 145
[25] Cox J C, Ingersoll J E and Ross S A 1985 A theory of the term structure of interest rates *Econometrica* **53** 385

[26] Johnson H and Shanno D 1987 Option pricing when the variance is changing *J. Financ. Quant. Anal.* **22** 143

[27] Kloeden P E and Platen E 2013 *Numerical Solution of Stochastic Differential Equations* (Berlin: Springer)

[28] Heston S L 1993 A closed-form solution for options with stochastic volatility with applications to bond and currency options *Rev. Financ. Stud.* **6** 327

[29] Hull J and White A 1987 The pricing of options on assets with stochastic volatilities *J. Financ.* **42** 281

[30] Bakshi G, Cao C and Chen Z 1997 Empirical performance of alternative option pricing models *J. Financ.* **52** 2003

[31] Shreve S 2004 *Stochastic Calculus for Finance II: Continuous-Time Models* (New York: Springer)

[32] Mandelbrot B B 1997 *Fractals and Scaling in Finance. Selecta Volume E* (New York: Springer)

[33] Liu Y, Gopikrishnan P, Cizeau P, Meyer M, Peng C-K and Stanley H E 1999 The statistical properties of the volatility of price fluctuations *Phys. Rev.* E **60** 1390

[34] Gopikrishnan P, Plerou V, Amaral L A N, Meyer M and Stanley H E 1999 Scaling of the distribution of fluctuations of financial market indices *Phys. Rev.* E **60** 5305

[35] Muzy J F, Delour J and Bacry E 2000 Modelling fluctuations of financial time series: from cascade process to stochastic volatility nodel *Eur. Phys. J.* B **17** 537

[36] Barndorff-Nielsen O E and Shephard N 2001 Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics *J. R. Stat. Soc.* B **63** 167

[37] Dragulescu A A and Yakovenko V M 2002 Probability distribution of returns in the Heston model with stochastic volatility *Quant. Financ.* **2** 443

[38] Seemann L, McCauley J L and Gunaratne G H 2011 Intraday volatility and scaling in high frequency foreign exchange markets *Int. Rev. Financ. Anal.* **20** 121

[39] Blanca P, Chicheportichea R and Bouchaud J-P 2014 The fine structure of volatility feedback: II. Overnight and intra-day effects *Physica* A **402** 58

[40] Feigenbaum J 2003 Financial physics *Rep. Prog. Phys.* **66** 1611

[41] Magdziarz M 2009 Black–Scholes formula in subdiffusive regime *J. Stat. Phys.* **136** 553

[42] Orzel S and Weron A 2010 Calibration of the subdiffusive Black–Scholes model *Acta Phys. Polon.* B **41** 1151

[43] Pirjol D, Jafarpour F and Iyer-Biswas S 2017 Phenomenology of stochastic exponential growth arXiv:1702.08035

[44] Mazzon A 2012 Processo square root *PhD Thesis* Milan University

[45] Oshanin G and Scher G 2012 Two stock options at the races: Black–Scholes forecasts *Quant. Financ.* **12** 1325

[46] Hu Y and Oksendal B 2003 Fractional white noise calculus and applications to finance *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **06** 1

[47] Biagini F, Hu Y, Øksendal B and Zhang T 2008 *Stochastic Calculus for Fractional Brownian Motion and Applications* (London: Springer) Ch 7

[48] Conus D and Wildman M 2016 A Gaussian Markov alternative to fractional Brownian motion for pricing financial derivatives arXiv:1608.03428

[49] Barndorff-Nielsen O E and Shephard N 2001 Modelling by Lévy processes for financial econometrics *Lévy Processes: Theory and Applications* ed O E Barndorff-Nielsen *et al* (London: Springer) pp 283–318

[50] Bates D S 1996 Jumps and stochastic volatility: exchange rate processes implicit in Deutsche Mark options *Rev. Financ. Stud.* **9** 69

[51] Peters O 2011 Optimal leverage from non-ergodicity *Quant. Financ.* **11** 1593

[52] Peters O and Klein W 2013 Ergodicity breaking in geometric Brownian motion *Phys. Rev. Lett.* **110** 100603

[53] Caravelli F, Sindoni L, Caccioli F and Ududec C 2015 Optimal leverage trajectories in presence of market impact *Phys. Rev.* E **94** 022315

[54] Hasbrouck J 2007 *Empirical Market Microstructure* (Oxford: Oxford University Press)

[55] Metzler R, Jeon J-H, Cherstvy A G and Barkai E 2014 Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking *Phys. Chem. Chem. Phys.* **16** 24128

[56] Nørregaard K, Metzler R, Ritter C, Berg-Sørensen K and Oddershede L 2017 Manipulation and motion of organelles and single molecules in living cells *Chem. Rev.* **117** 4342

[57] Montroll E W and Schlesinger M F 1982 On $1/f$ noise and other distributions with long tails *Proc. Natl Acad. Sci. USA* **79** 3380

[58] Schulz J H P, Barkai E and Metzler R 2014 Aging renewal theory and application to random walks *Phys. Rev.* X **4** 011028

[59] Chubynsky M V and Slater G W 2014 Diffusing diffusivity: a model for anomalous, yet Brownian, diffusion *Phys. Rev. Lett.* **113** 098302

[60] Jain R and Sebastian K L 2016 Diffusing diffusivity: survival in a crowded rearranging and bounded domain *J. Phys. Chem.* B **120** 3988

[61] Chechkin A V, Seno F, Metzler R and Sokolov I M 2017 Brownian yet non-Gaussian diffusion: from superstatistics to subordination of diffusing diffusivities *Phys. Rev.* X **7** 021002

[62] Renner C, Peinke J and Friedrich R 2001 Evidence of Markov properties of high frequency exchange rate data *Physica* A **298** 499

[63] Allez R and Bouchaud J-P 2011 Individual and collective stock dynamics: intra-day seasonalities *New J. Phys.* **13** 025010

[64] Kleczkowski A, Oles K, Gudowska-Nowak E and Gilligan C A 2012 Searching for the most cost-effective strategy for controlling epidemics spreading on regular and small-world networks *J. R. Soc. Interface* **9** 158

[65] Forster G A and Gilligan C A 2007 Optimizing the control of disease infestations at the landscape scale *Proc. Natl Acad. Sci. USA* **104** 4984

[66] Iyer-Biswas S *et al* 2014 Scaling laws governing stochastic growth and division of single bacterial cells *Proc. Natl Acad. Sci. USA* **111** 15912