

Directed continuous-time random walk with memory

Jaroslav Klamut^{1,a} and Tomasz Gubiec^{1,2}

¹ Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Pasteur Str. 5, 02093 Warsaw, Poland

² Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, USA

Received 11 July 2018 / Received in final form 8 January 2019

Published online 3 April 2019

© The Author(s) 2019. This article is published with open access at [Springerlink.com](https://www.springerlink.com)

Abstract. In this paper, we are addressing the old problem of long-term nonlinear autocorrelation function versus short-term linear autocorrelation function. As continuous-time random walk (CTRW) can describe almost all possible kinds of diffusion, it seems to be an excellent tool to use. To be more precise, for instance, CTRW can successfully describe the short-term negative autocorrelation of returns in high-frequency financial data (caused by the bid-ask bounce phenomena). We observe long-term autocorrelation of absolute values of returns. Can it also be described by the CTRW model? And maybe more importantly, to what extent can it be explained by the same phenomena? To refer to these questions, we propose a new directed CTRW model with memory. The canonical CTRW trajectory consists of spatial jumps preceded by waiting times. In directed CTRW, we consider the case with positive spatial jumps only. We take into account the memory in the model as each spatial jump depends on the previous one. This model, based on simple assumptions, allowed us to obtain the general formula covering most popular types of nonlinear autocorrelation functions.

1 Introduction

In 1956 two physicists Montroll and Weiss, in the context of dispersive transport diffusion, introduced a new stochastic process they named continuous-time random walk (CTRW) [1]. As dynamics of many complex systems can be described by discrete spatiotemporal events, i.e., the spatial jump of stochastic process preceded by waiting time, the formalism of CTRW seems a natural description. On the other hand, CTRW can be considered as a way to introduce finite, continuous and fluctuating interevent times into a random walk. Since its introduction, the concept of CTRW found many applications and extensions [2].

The CTRW has initially described a photocurrent relaxation in amorphous films [3–7]. A broader spectrum of other applications and arrangements contains: diffusion in probabilistic fractal structures (percolations clusters [8] and fractal diffusion [9]), aging of glasses [10,11], nearly constant dielectric loss in disordered ionic conductors [12], cardiological rhythms [13], electron transfer [14], search models [15], transport in porous media [16], diffusion of epicenters of earthquake aftershocks [17], subsurface tracer diffusion [18], hydrogen diffusion in nanostructure compounds [19] or even human travel [20]. In this work, we are particularly interested in CTRW models used in the description of financial markets, mainly financial time series, where the dependencies and distributions of times

between transactions and price changes are considered [21–36].

In the majority of cases, the analyzed CTRW models focus on the spatial distribution with zero mean or even symmetric distribution. In other words, the drift term is usually neglected. The case of drift was studied in [37] (and references therein). The case of canonical CTRW, where both spatial and temporal distributions are i.i.d. and they do not depend on each other, turns out to be a compelling model, able to describe many cases of normal or anomalous diffusion. Different types of CTRW are obtained if the mean waiting time is finite or diverging (but assuming finite variance of the spatial distribution). In the first case, we observe a normal diffusion, in the latter subdiffusion occurs [38]. If the variance of the spatial distribution diverges and waiting time distribution (WTD) has a finite mean, we obtain the description of Lévy flights.

The other promising branch of CTRW models is the one considering memory, i.e., the dependence between successive jumps. Different types of dependencies were already studied: the backward or forward correlations between spatial jump directions [39] in the case of concentrated lattice gas for the study of the tracer coefficient [40], even taking into account the dependencies over several subsequent jumps [41]. Also, models driven by the negative feedback in consecutive jumps were built, considering one-step memory [35,42] and later two-step or even infinite-step memory [34]. Their potential applications cover the Le Chatelier–Braun principle of contrariness.

^a e-mail: jaroslav.klamut@fuw.edu.pl

Memory in waiting times also appeared in some CTRW models [43–48]. Examples of used dependencies are correlations which solely depend on the sign of consecutive jumps [43], random walk of waiting times [45,46], exponential and slowly decaying persistent power-law correlations [47].

Our work is directly motivated by the application of CTRW in the description of high-frequency financial data. The universal properties of all financial price time series are sometimes referred to as stylized facts [49,50]. There are two well known stylized facts about autocorrelation of price time series. The first one states that the time-dependent autocorrelation of price increments (or logarithmic returns) is negative and quickly decays to zero [43]. A CTRW model with memory [35] successfully reproduced this fact. The second stylized fact states that autocorrelation of the absolute value of price increments (or absolute values of log returns) is a positive slowly decaying function. Also, the amplitude in the second case is usually an order of magnitude higher than in the first case. It is the reminiscence of the so-called volatility clustering phenomenon [51]. It seems natural to ask if the CTRW model with memory introduced in [35] adapted to describe absolute values of price changes can successfully reproduce the second mentioned stylized fact. We are answering this question below.

The paper is organized as follows: in Section 2, we present the motivation behind our work and define and solve the proper stochastic process. In Section 3, we obtain velocity autocorrelation function (VAF) and in Section 4 the comparison with empirical data is made. The intraday-seasonality is taken into account in Section 5. Finally in Section 6, we conclude with remarks on the results presented in this paper.

2 Model

The CTRW trajectory process is a step function (a piecewise constant function) that can be interpreted as spatial jumps preceded by waiting times. In canonical CTRW, all jumps and waiting times are i.i.d. random variables. The model that successfully described the short term linear autocorrelation of the stock market returns was based on the dependence of two consecutive price changes in the following form

$$h(r_n|r_{n-1}) = (1 - \epsilon)h(r_n) + \epsilon\delta(r_n + r_{n-1}), \quad (1)$$

where $h(r_n|r_{n-1})$ is the probability density of the price change r_n under the condition of the previous price change value r_{n-1} . This assumption was based on the technical property of the double auction market named bid-ask bounce. The CTRW formalism allows calculating linear autocorrelation function analytically. It is worth to mention that by changing the distribution h on the right-hand side of (1) we can obtain some nonlinear autocorrelation functions, but only in the limited cases of monotonic nonlinearity. Obtaining the autocorrelation of absolute values of returns cannot be achieved this way. To do so, we construct the directed continuous-time random

walk (DCTRW) focused on the absolute values of spatial jumps. This process is defined in such a way that its linear autocorrelation is exactly the nonlinear autocorrelation of absolute values of returns. As in (1) we consider one-step memory for consecutive jumps and no dependence between waiting times or between waiting times and jumps. We take modules of price changes $R_n = |r_n|$ and obtain dependence in the form

$$H(R_n|R_{n-1}) = (1 - \epsilon)H(R_n) + \epsilon\delta(R_n - R_{n-1}), \quad (2)$$

where $H(R_n)$ and $H(R_n|R_{n-1})$ are respectively distribution of jump modules and conditional distribution of jump modules. Parameter ϵ describes the strength of the memory, for $\epsilon = 0$ we obtain the model without memory. Considering DCTRW of absolute values of price changes, Dirac delta describes the same consecutive jumps, not the opposite ones, as it was the case in (1). As in the previous case, other types of the nonlinearity can be obtained by changing the distribution H on the right-hand side of (2). To sum up, our model can be described by the probability density functions of n th jump R_n after waiting time t_n conditioned on all previous R_i and t_i :

$$\rho(R_n, t_n|R_{n-1}, t_{n-1}; \dots; R_1, t_1) = H(R_n|R_{n-1})\psi(t), \quad (3)$$

where $\psi(t)$ represents the WTD. Results will be presented for any WTD and in two specific cases.

To make the process stationary, we cannot use the same WTD for the first jump as well as for other jumps [52,53]. This is because the previous (preinitial) jump might have occurred at any time before $t = 0$. Therefore, we should define

$$\psi_1(t) = \frac{\int_0^\infty dt' \psi(t+t')}{\int_0^\infty dt'' \int_0^\infty dt' \psi(t'+t'')}, \quad (4)$$

as the WTD before the first jump. Moreover, for simplicity of notation it is useful to introduce sojourn probability $\Psi(t) = \int_t^\infty \psi(t')dt'$. Above probabilities can be easily expressed in the Laplace domain:

$$\begin{aligned} \tilde{\psi}(s) &= \mathcal{L}[\psi(t)], \\ \tilde{\Psi}(s) &= \frac{1 - \tilde{\psi}(s)}{s}, \\ \tilde{\psi}_1(s) &= \frac{1 - \tilde{\psi}(s)}{\langle t \rangle s}, \\ \tilde{\Psi}_1(s) &= \frac{1 - \tilde{\psi}_1(s)}{s}, \end{aligned} \quad (5)$$

where $\mathcal{L}[\cdot]$ denotes Laplace transform and $\langle t \rangle = \int_0^\infty t\psi(t)dt < \infty$ is expected (mean) waiting time. On our path to obtaining nonlinear VAF of the CTRW by calculating the linear VAF of DCTRW we need a few intermediate quantities. One of these intermediate dynamic quantities describing the stochastic process is the stochastic, sharp, n -step propagator $Q_n(X, R_n; t|\xi)$, $n = 1, 2, \dots$. This propagator is defined as the conditional probability density that the price,

which was initially (at $t = 0$) in the origin value ($X = 0$) reached by preinitial jump ξ , makes its n th jump by R_n from $X - R_n$ to X exactly at time t . $\tilde{Q}_n(K, R_n; s|\xi)$ is sharp propagator in the Fourier–Laplace domain. The recurrence relation between two successive sharp stochastic propagators can be written for any form of $H(R_n|R_{n-1})$ and $n > 1$, as follows:

$$\begin{aligned} \tilde{Q}_n(K, R_n; s|\xi) &= \tilde{\psi}(s)e^{iKR_n} \int_{-\infty}^{\infty} dR_{n-1} H(R_n|R_{n-1}) \tilde{Q}_{n-1}(K, R_{n-1}; s|\xi). \end{aligned} \tag{6}$$

The first sharp propagator $Q_1(X, R_1; t|\xi)$ can be calculated directly from definition

$$\begin{aligned} Q_1(X, R_1; t|\xi) &= Q_1(X; t|\xi)\delta(X - R_1) \\ &= \psi_1(t)H(X|\xi)\delta(X - R_1). \end{aligned} \tag{7}$$

The following sharp propagators can be calculated using equation (6). After integrating over R_n we obtain the recurrence relation

$$\begin{aligned} \frac{\tilde{Q}_n(K; s|\xi)}{\tilde{\psi}(s)} &= (1 - \epsilon)\tilde{H}(K)\tilde{Q}_{n-1}(K; s|\xi) \\ &+ \epsilon \int_0^{\infty} dR_{n-1} e^{iKR_{n-1}} \tilde{Q}_{n-1}(K, R_{n-1}; s|\xi). \end{aligned} \tag{8}$$

The second intermediate quantity is the soft propagator $P(X, t)$, defined as the conditional probability density that the price, which was initially (at $t = 0$) in the origin value ($X = 0$) reached by preinitial jump ξ , at time t is equal to X . We can write the relation between the soft propagator $P(X, t)$ and the sharp propagator $Q(x, t)$ (in the Fourier–Laplace domain) as

$$\tilde{P}(K; s) = \tilde{\Psi}_1(s) + \tilde{\Psi}(s)\tilde{Q}(K; s), \tag{9}$$

$$\tilde{Q}(K; s) = \sum_{n=1}^{\infty} \tilde{Q}_n(K; s). \tag{10}$$

To obtain explicit formula for the right hand side of equation (10), in case of one-step memory defined by equation (2), we use the recurrence relation (8). The precise description of this procedure can be found in Appendix A. The result can be simply substituted into equation (9) and hence we obtain the soft propagator in the following form

$$\begin{aligned} \tilde{P}(K; s) &= \frac{1}{s} - \frac{1 - \tilde{\psi}(s)}{\langle t \rangle s^2} \\ &+ \frac{[1 - \tilde{\psi}(s)]^2}{\langle t \rangle s^2} \frac{S(K; s)}{1 - (1 - \epsilon)\tilde{\psi}(s)S(K; s)}, \end{aligned} \tag{11}$$

where

$$S(K; s) = \sum_{n=1}^{\infty} (\tilde{\psi}(s)\epsilon)^{n-1} \tilde{H}(nK). \tag{12}$$

As a result, the soft propagator in the Fourier–Laplace domain takes a reasonably simple form; however, it still contains the function S which is given as an infinite sum. Fortunately, to compute moments of the process and VAF we need to know the corresponding derivatives of the soft propagator at point $K = 0$, which can be determined explicitly.

The first and the second moment of the process in the Fourier–Laplace domain are

$$\tilde{m}_1(s) = -i \left. \frac{\partial \tilde{P}(K; s)}{\partial K} \right|_{K=0} = \frac{M_1}{\langle t \rangle s^2}, \tag{13}$$

$$\begin{aligned} \tilde{m}_2(s) &= - \left. \frac{\partial^2 \tilde{P}(K; s)}{\partial K^2} \right|_{K=0} \\ &= \frac{M_2 + (1 - \epsilon)(2M_1^2 - M_2)\tilde{\psi}(s) - \epsilon M_2 \tilde{\psi}^2(s)}{\langle t \rangle s^2 (1 - \tilde{\psi}(s))(1 - \epsilon \tilde{\psi}(s))} \\ &= \frac{M_2(1 + \epsilon \tilde{\psi}(s))}{\langle t \rangle s^2 (1 - \epsilon \tilde{\psi}(s))} + \frac{2(1 - \epsilon)\tilde{\psi}(s)M_1^2}{\langle t \rangle s^2 (1 - \tilde{\psi}(s))(1 - \epsilon \tilde{\psi}(s))}, \end{aligned} \tag{14}$$

where M_i is the i th moment of jump modules distribution $H(R)$. The first moment of the directed process in the time space rises linearly in time $m_1(t) = \frac{M_1}{\langle t \rangle} t$, exactly like for the process without memory. It is worth to notice that it does not depend on ϵ .

3 Nonlinear velocity autocorrelation function of CTRW

In the general case, the VAF in the time domain is given by

$$C(t) = \frac{1}{2} \ddot{m}_2(t) - \dot{m}_1^2(t). \tag{15}$$

In the case of the DCTRW process considered in this paper, it takes the form

$$\begin{aligned} C(t) &= \left(\frac{M_2 - M_1^2}{2 \langle t \rangle} \right) \mathcal{L}^{-1} \left[\frac{1 + \epsilon \tilde{\psi}(s)}{1 - \epsilon \tilde{\psi}(s)} \right] \\ &+ \frac{M_1^2}{2 \langle t \rangle} \mathcal{L}^{-1} \left[\frac{1 + \tilde{\psi}(s)}{1 - \tilde{\psi}(s)} - \frac{2}{\langle t \rangle s} \right], \end{aligned} \tag{16}$$

where $\mathcal{L}^{-1}[\cdot]$ is the inverse Laplace transform. To investigate the behavior of our nonlinear VAF in the limits $t \rightarrow 0$ and $t \rightarrow \infty$, we have to check the behavior in limits $s \rightarrow \infty$ and $s \rightarrow 0$ of the expressions inside inverse Laplace transforms. It is known that for $s \rightarrow \infty$, $\tilde{\psi}(s)$ goes to 0, while for $s \rightarrow 0$ the approximation $\tilde{\psi}(s) \approx 1 - \langle t \rangle s$ should be

used. Therefore, in the limit of long times VAF vanishes and for short times we obtain the variance of the process $C(t) \approx \frac{M_2}{2\langle t \rangle} \delta(t)$. Normalized VAF $C^n(t)$ has Dirac delta at $t = 0$ so $C^n(t) = \frac{2\langle t \rangle}{M_2} C(t)$.

To compare our model with empirical data we use two specific WTDs: exponential and double-exponential, with explicit results for both. First one is a simple distribution and its characteristics match stylized facts of financial time series. The second one can be fitted to empirical data with high accuracy and still allows to obtain analytical VAF from the model. Exponential WTD with the mean waiting time equals to $\langle t \rangle$ is given as

$$\psi(t) = \frac{1}{\langle t \rangle} \exp\left(-\frac{t}{\langle t \rangle}\right), \quad (17)$$

and double-exponential WTD with partial mean waiting times equals to τ_1 and τ_2 and weighting parameter w is given as

$$\psi(t) = \frac{w}{\tau_1} \exp\left(-\frac{t}{\tau_1}\right) + \frac{1-w}{\tau_2} \exp\left(-\frac{t}{\tau_2}\right). \quad (18)$$

Mean waiting time of double-exponential WTD is $\langle t \rangle = w\tau_1 + (1-w)\tau_2$.

For the exponential WTD, nonlinear VAF is easily expressed as

$$C^n(t) = \delta(t) + \frac{2\epsilon(1-M)}{\langle t \rangle} \exp\left(-\frac{(1-\epsilon)t}{\langle t \rangle}\right), \quad (19)$$

where

$$M = \frac{M_1^2}{M_2} \in (0; 1). \quad (20)$$

Although exponential WTD does not describe properly the empirical WTD, one can easily interpret meaning of the parameters. Firstly, for $t > 0$ nonlinear VAF is positive (unlike the negative linear VAF) and decreases exponentially. Relaxation time increases and the amplitude reduces with longer mean waiting time. Increasing parameter ϵ results in higher relaxation time and amplitude, especially for $\epsilon = 0$ VAF is non-zero only for $t = 0$. It is also noticeable that normalized nonlinear VAF depends only on a ratio between the first moment squared and the second moment of jumps modules. The amplitude of nonlinear VAF decreases with the increase of M .

For the double-exponential WTD, which satisfactorily fits empirical data, normalized VAF is

$$\begin{aligned} C^n(t) &= \delta(t) + A_0 e^{-v_0 t} + A_1 e^{-v_1 t} + A_2 e^{-v_2 t}, \\ w_i &= \tau_i^{-1}, \\ v &= w w_1 + (1-w) w_2, \\ v_0 &= (1-w) w_1 + w w_2, \\ v_i &= \frac{1}{2} \left[w_1 + w_2 - \epsilon v - (-1)^i \right. \\ &\quad \left. \times \sqrt{(w_1 + w_2 - \epsilon v)^2 - 4w_1 w_2 (1 - \epsilon)} \right], \end{aligned}$$

$$\begin{aligned} A_0 &= 2 \frac{M}{v_0} w(1-w)(w_1 - w_2)^2, \\ A_i &= (-1)^i \frac{2\epsilon(1-M)}{v_1 - v_2} [w_1 w_2 - v v_i], \\ i &\in \{1, 2\}. \end{aligned} \quad (21)$$

There are three exponential functions in this formula, except Dirac delta, all with positive amplitudes. The first one is worth mentioning, as it does not depend on ϵ . It implicates that for DCTRW processes linear VAF (equal to nonlinear VAF for CTRW) can be non-zero even in the case without memory ($\epsilon = 0$). That effect does not occur in linear VAF of the original CTRW. In the limit when two-exponent WTD goes to exponential WTD ($w \rightarrow 0$ or 1) this element vanishes. Two other exponential functions depend on ϵ and describe decaying at different rates. Similarly like for exponential WTD, increasing ϵ results in higher VAF. In the case of spatial changes with zero variance ($M \rightarrow 1$), these terms are equal to zero.

4 Empirical results

To compare our model with empirical data, we use tick-by-tick transaction data from Polish stock market (Warsaw Stock Exchange) from years 2011–2012. Presented results are calculated for KGHM (copper mining company) – one of the most liquid stocks. We extract waiting times (periods between transactions) and jumps (price changes) from this data. Comparing the model with empirical data requires estimating parameters. We obtain τ_1 , τ_2 , w from fitting two-exponential WTD to empirical histogram using least squares method. We calculate two first moments of price changes absolute values M_1 and M_2 explicitly from the empirical distribution. Parameter ϵ is calculated as one-step autocorrelation of price changes absolute values. Empirical linear VAF and nonlinear VAF were calculated using the method described in [35]. We want to remind that linear VAF obtained from the CTRW model based on (1) (and its modifications in [34,54]) built using these parameters fits satisfactorily with empirical data. This result is presented in Figure 1, as linear VAF of CTRW (dashed line) with good approximation fits the empirical point (gray dots). The situation is completely different in the case of nonlinear VAF. As shown in Figure 1, theoretical approach does not explain observed nonlinear VAF (solid line, and black dots). Theoretical nonlinear VAF decays slower than linear VAF, but empirical nonlinear VAF decays much slower than both of them. There can be many reasons explaining obtained disagreement: daily seasonality, long-term jump modules dependencies and long-term waiting times dependencies (see Fig. 2). In the next section, we check if taking into account the first one (daily seasonality) can significantly improve the quality of the description of data. Taking into account the two latter cases require new CTRW models with memory and go beyond the scope of this article.

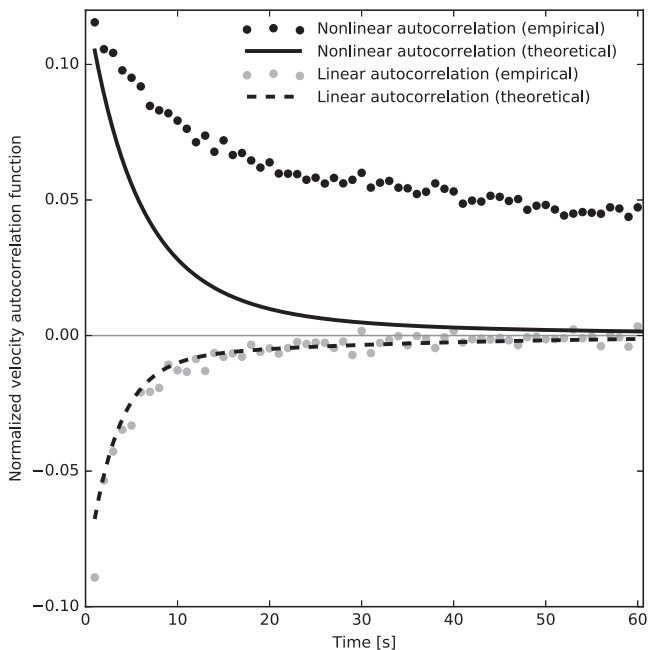


Fig. 1. Empirical, theoretical and linear, nonlinear VAF of KGHM price changes (one of the most liquid stock at Warsaw Stock Exchange) in years 2011–2012. Points are empirical result (gray for linear VAF and black for nonlinear), lines represent theoretical results (dashed for linear VAF, solid for nonlinear VAF (23)). As theoretical linear VAF fits the empirical counterpart, the nonlinear VAF does not fit empirical data. Fitted values of the parameters are: $M = 0.269$, $\epsilon = 0.258$, $\tau_1 = 3.63$, $\tau_2 = 32.57$, $w = 0.586$, $p = 14986$, $q = 2.25 \times 10^8$.

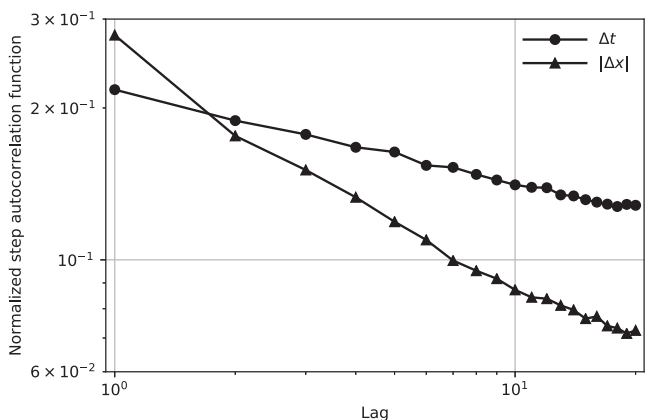


Fig. 2. Normalized step autocorrelation function of waiting times and the modules of price changes for KGHM in years 2011–2012 in the logarithmic coordinates.

5 Nonstationarity

Varying mean inter-trade time during trading session is a stylized fact observed on stock prices on every market [50,55,56]. This intraday pattern, often called the “lunch effect”, is characterized by low volatility and long inter-trade times in the middle of the day. On the contrary, activity on the stock market is higher at the beginning and the end of a trading session. This effect influences

both linear and nonlinear VAF of the process, as described in [54], where the general formula for the impact of this phenomenon is given. Following [54] we describe the daily pattern with the rational function

$$\theta(t) = \frac{1}{a[(t - p)^2 + q]}, \tag{22}$$

where $\theta(t)$ can be interpreted as varying mean of the waiting times distribution during a trading session. Now, we can obtain an explicit expression for normalized VAF taking the seasonality into account:

$$C^n(t) = \sum_{j=0}^2 A_j J v_j^{-\frac{1}{2}} e^{-v_j \tau_{\min}} \left[\text{erf}(\sqrt{v_j J_0}) + \text{erf}(\sqrt{v_j J_k}) \right],$$

$$J = \frac{1}{2(T - t)} \sqrt{\frac{\pi X}{t}},$$

$$J_0 = \frac{t}{X} \left(\frac{t}{2} - p \right)^2,$$

$$J_k = \frac{t}{X} \left(\frac{t}{2} - (p - T) \right)^2,$$

$$\tau_{\min} = \frac{t^2 + q}{X} t,$$

$$X = \frac{T^2}{3} - pT + p^2 + q, \tag{23}$$

where T is the length of a day, parameters p and q are fitted to data and come from the rational form of day seasonality and erf is the error function. Unfortunately taking nonstationarity into account only slightly improves the results, making a difference that is within the line width in Figure 1 (solid line). This result implies that definitely it is not the daily volatility pattern that creates the disagreement between empirical and theoretical nonlinear VAF.

6 Conclusions

We proposed a methodology of calculating the general nonlinear VAF within the CTRW model, by creating specifically suited DCTRW. The nonlinear VAF of the CTRW can be obtained as linear VAF of the DCTRW. In the specific case of CTRW created to describe stock price in time, we solved DCTRW model with one-step memory in jumps and obtained the analytical equation for propagator, first two moments and VAF (21). The obtained nonlinear VAF shows interesting properties: it is positive; it decays exponentially but slower than the linear VAF; nonlinear VAF can be non-zero even without any memory ($\epsilon = 0$) for non-exponential WTD. Next, we considered nonstationarity in the form of rational function (22) and obtained analytical nonlinear VAF (23).

The presented simple model turned out to be unable to describe empirical data (Fig. 1). This result implies that simple bid-ask bounce phenomenon is not sufficient

to explain long memory in financial time series and volatility clustering. Moreover, taking into account the daily seasonality of the time series does not change this result.

The CTRW framework with memory in price increments seems to be a perfect tool to describe stock market price changes, but as our results suggest a different approach is necessary. In Figure 2, we present empirical step autocorrelation functions of waiting times and modules of price changes. Focusing on the correlation of two consecutive modules of price changes and two consecutive waiting times may suggest that the memory in price changes is more important and the dependence in waiting times can be neglected. Looking at the whole plot suggests exactly the opposite, as the autocorrelation in waiting times decays much slower than the autocorrelation of modules of price changes. We suggest that taking into account the long memory in waiting times would improve the results, but it is still a great challenge and an open problem.

Author contribution statement

T. Gubiec set the idea for this paper, contributed to the calculations and discussed the results. J. Klamut did the calculations, data analysis and graphs. Both authors participated in the writing of the manuscript.

Open Access This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Appendix A: Closed-form of the soft propagator $\tilde{P}(K; s)$

In this Appendix, we present the details of derivation of the soft propagator $\tilde{P}(K; s)$ (11). Firstly, we recursively expand the integral in relation (8) using equations (6) and (7). We obtain the recurrence relation, expressing each n -step sharp propagators in terms of all sharp propagators with all smaller n .

$$\begin{aligned} \tilde{Q}_n(K; s|\xi) &= \tilde{\psi}(s)(1 - \epsilon) \left(\sum_{j=1}^{n-1} [\tilde{\psi}(s)\epsilon]^{j-1} \tilde{H}(jK) \tilde{Q}_{n-j}(K; s|\xi) \right) \\ &+ [\tilde{\psi}(s)\epsilon]^{n-1} \tilde{\psi}(s) [(1 - \epsilon)\tilde{H}(nK) + \epsilon e^{niK\xi}], \quad n \geq 2. \end{aligned} \tag{A.1}$$

The one-step propagator can be calculated directly from its definition

$$\tilde{Q}_1(K; s|\xi) = \tilde{\psi}_1(s) [(1 - \epsilon)\tilde{H}(K) + \epsilon e^{iK\xi}]. \tag{A.2}$$

As the type of recurrence relation is completely different than in work [35], we present a different method to solve

this equation. Firstly, we do not need to know exactly each n -step propagator, our goal is to obtain the general propagator given by equation (10). By reordering and grouping the same step sharp propagators in summation of all $\tilde{Q}_n(K; s|\xi)$, one can calculate the general propagator:

$$\begin{aligned} \tilde{Q}(K; s|\xi) &= \tilde{\psi}_1(s) \left[(1 - \epsilon)S(K; s) + \epsilon \sum_{n=1}^{\infty} [\tilde{\psi}(s)\epsilon]^{n-1} e^{niK\xi} \right] \\ &+ \tilde{\psi}(s)(1 - \epsilon)S(K; s)\tilde{Q}(K; s|\xi), \end{aligned} \tag{A.3}$$

where we introduce $S(K; s) = \sum_{n=1}^{\infty} [\tilde{\psi}(s)\epsilon]^{n-1} \tilde{H}(nK)$. This gives us closed-form of the sharp propagator

$$\begin{aligned} \tilde{Q}(K; s|\xi) &= \frac{\tilde{\psi}_1(s)(1 - \epsilon)S(K; s) + \tilde{\psi}_1(s)\epsilon \sum_{n=1}^{\infty} [\tilde{\psi}(s)\epsilon]^{n-1} e^{niK\xi}}{1 - \tilde{\psi}(s)(1 - \epsilon)S(K; s)}. \end{aligned} \tag{A.4}$$

Now, $\tilde{Q}(K; s)$ can be calculated by integrating $\tilde{Q}(K; s|\xi)$ over all possible pre-initial jumps ξ , which come from distribution $H(\xi)$

$$\tilde{Q}(K; s) = \frac{\tilde{\psi}_1(s)S(K; s)}{1 - (1 - \epsilon)\tilde{\psi}(s)S(K; s)}. \tag{A.5}$$

From that we can obtain the soft propagator using equation (9).

Appendix B: Intraday seasonality

In this Appendix, we explain more precisely how daily seasonality can be considered. According to [54], we assume that real nonstationary process $Y(t)$ can be expressed as a stationary process $X(f(t))$ with its own time. $\tau = f(t)$ is a transformation of real time, that ensures considered equality between processes X and Y . The model proposed in this work describes stationary process X , we would like to see what is autocorrelation C_Y of nonstationary process Y knowing the autocorrelation C_X of X . In our case, it means that following relation must hold:

$$\frac{dt}{\theta(t)} = \frac{\tau}{\langle t \rangle}. \tag{B.1}$$

Then, the autocorrelation of Y is given by

$$\begin{aligned} C_Y(t) &= \int_{\tau_{\min}}^{\tau_k} WC_X(\tau) d\tau + \int_{\tau_{\min}}^{\tau_0} WC_X(\tau) d\tau, \\ \tau_0 &= \left(\frac{t^2}{3} - pt + p^2 + q \right) \frac{t}{X}, \\ \tau_k &= \left(\frac{t^2}{3} - (p - T)t + (p - T)^2 + q \right) \frac{t}{X}, \\ W &= X \left(2t(T - t) \sqrt{X \frac{\tau}{t} - q - \frac{t^2}{12}} \right)^{-1} \end{aligned} \tag{B.2}$$

and the rest of the variables are defined in main text. To obtain the final result given by equation (23) we have to calculate integrals above.

References

1. E.W. Montroll, G.H. Weiss, *J. Math. Phys.* **6**, 167 (1965)
2. R. Kutner, J. Masoliver, *Eur. Phys. J. B* **90**, 50 (2017)
3. H. Scher, E.W. Montroll, *Phys. Rev. B* **12**, 2455 (1975)
4. G. Pfister, H. Scher, *Adv. Phys.* **27**, 747 (1978)
5. E.W. Montroll, M.F. Schlesinger, in *Nonequilibrium Phenomena II: From Stochastics to Hydrodynamics*, edited by J. Lebowitz, E. Montroll (North-Holland, Amsterdam, 1984), pp. 1–121
6. G. Weiss, in *Fractals in Science* (Springer, Berlin, 1994), pp. 119–162
7. J.P. Bouchaud, A. Georges, *Phys. Rep.* **195**, 127 (1990)
8. D. Ben-Avraham, S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge University Press, Cambridge, 2000)
9. R. Hilfer, *Physica A* **329**, 35 (2003)
10. E. Barkai, Y.C. Cheng, *J. Chem. Phys.* **118**, 6167 (2003)
11. C. Monthus, J.P. Bouchaud, *J. Phys. A, Math. Gen.* **29**, 3847 (1996)
12. W. Dieterich, P. Maass, *Solid State Ion.* **180**, 446 (2009)
13. N. Iyengar, C. Peng, R. Morin, A. Goldberger, L. Lipsitz, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **271**, R1078 (1996)
14. J. Nelson, *Phys. Rev. B* **59**, 15374 (1999)
15. M. Lomholt, K. Tal, R. Metzler, J. Klafter, *Proc. Natl. Acad. Sci.* **105**, 11055 (2008)
16. G. Margolin, B. Berkowitz, *J. Phys. Chem. B* **104**, 3942 (2000)
17. A. Helmstetter, D. Sornette, *Phys. Rev. E* **66**, 061104 (2002)
18. H. Scher, G. Margolin, R. Metzler, J. Klafter, B. Berkowitz, *Geophys. Res. Lett.* **29**, 5 (2002)
19. R. Hempelmann, in *Anomalous Diffusion From Basics to Applications*, edited by R. Kutner, A. Pękalski, K. Sznajd-Weron (Springer, Berlin, Heidelberg, 1999), pp. 247–252
20. L. Hufnagel, D. Brockmann, T. Geisel, *Nature* **439**, 462 (2006)
21. E. Scalas, R. Gorenflo, F. Mainardi, *Physica A* **284**, 376 (2000)
22. F. Mainardi, M. Raberto, R. Gorenflo, E. Scalas, *Physica A* **287**, 468 (2000)
23. M. Raberto, E. Scalas, F. Mainardi, *Physica A* **314**, 749 (2002)
24. E. Scalas, R. Gorenflo, F. Mainardi, *Phys. Rev. E* **69**, 011107 (2004)
25. E. Scalas, *Physica A* **362**, 225 (2006)
26. R. Kutner, F. Świtłała, *Quant. Financ.* **3**, 201 (2003)
27. J. Masoliver, M. Montero, G.H. Weiss, *Phys. Rev. E* **67**, 021112 (2003)
28. P. Repetowicz, P. Richmond, *Physica A* **344**, 108 (2004). (Applications of Physics in Financial Analysis 4 (APFA4))
29. J. Masoliver, M. Montero, J. Perelló, G.H. Weiss, *J. Econ. Behav. Organ.* **61**, 577 (2006)
30. J. Masoliver, M. Montero, J. Perello, G.H. Weiss, *Physica A* **379**, 151 (2007)
31. R. Kutner, *Phys. A, Stat. Mech. Appl.* **314**, 786 (2002)
32. E. Scalas, in *The Complex Networks of Economic Interactions* (Springer, Berlin, Heidelberg, 2006), pp. 3–16
33. J. Perelló, J. Masoliver, A. Kasprzak, R. Kutner, *Phys. Rev. E* **78**, 036108 (2008)
34. T. Gubiec, R. Kutner, *Eur. Phys. J. B* **90**, 228 (2017)
35. T. Gubiec, R. Kutner, *Phys. Rev. E* **82**, 046119 (2010)
36. A. Kasprzak, R. Kutner, J. Perelló, J. Masoliver, *Eur. Phys. J. B* **76**, 513 (2010)
37. R. Metzler, J. Klafter, *Phys. Rep.* **339**, 1 (2000)
38. H. Scher, M.F. Shlesinger, J.T. Bendler, *Phys. Today* **44**, 26 (1991)
39. J.W. Haus, K.W. Kehr, *Phys. Rep.* **150**, 263 (1987)
40. K. Kehr, R. Kutner, K. Binder, *Phys. Rev. B* **23**, 4931 (1981)
41. R. Kutner, *J. Phys. C: Solid State Phys.* **18**, 6323 (1985)
42. T. Gubiec, R. Kutner, *Acta Phys. Pol. A* **117**, 669 (2010)
43. M. Montero, J. Masoliver, *Phys. Rev. E* **76**, 061115 (2007)
44. M. Montero Torralbo, *Phys. Rev. E* **84**, 051139 (2011)
45. V. Tejedor, R. Metzler, *J. Phys. A: Math. Theor.* **43**, 082002 (2010)
46. M. Magdziarz, R. Metzler, W. Szcotoka, P. Zebrowski, *Phys. Rev. E* **85**, 051103 (2012)
47. A.V. Chechkin, M. Hofmann, I.M. Sokolov, *Phys. Rev. E* **80**, 031112 (2009)
48. J.H. Jeon, N. Leijnse, L.B. Oddershede, R. Metzler, *New J. Phys.* **15**, 045011 (2013)
49. R. Tsay, *Analysis of Financial Time Series*, 2nd edn., Wiley Series in Probability and Statistics (Wiley-Interscience, Hoboken, NJ, 2005)
50. R. Cont, *Quant. Financ.* **1**, 223 (2001)
51. R. Cont, in *Fractals in Engineering*, edited by J. Lévy-Véhel, E. Lutton (Springer, London, 2005), pp. 159–179
52. J.K.E. Tunaley, *Phys. Rev. Lett.* **33**, 1037 (1974)
53. J.K.E. Tunaley, *J. Stat. Phys.* **11**, 397 (1974)
54. T. Gubiec, M. Wiliński, *Physica A* **432**, 216 (2015)
55. J. Hasbrouck, *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading* (Oxford University Press, Oxford, 2007)
56. M.M. Dacorogna, R. Gencay, U. Muller, R.B. Olsen, O.V. Pictet, *An Introduction to High Frequency Finance* (Academic Press, New York, 2001)