'Continuous' Time Random Walk in Continuous Time Random Walk.

The crucial role of inter-event times in volatility clustering

Jaroslaw Klamut and Tomasz Gubiec

Institute of Experimental Physics
Faculty of Physics
University of Warsaw
(Dated: September 12, 2019)

Abstract

We are introducing the new family of the Continuous Time Random Walks (CTRW) with long-term memory within consecutive waiting times. This memory is introduced to the model by the assumption that consecutive waiting times are the analog of CTRW themselves. This way, we obtain the 'Continuous' Time Random Walk in Continuous Time Random Walk. Surprisingly, this type of process, only with the long-term memory within waiting times, can successfully describe slowly decaying nonlinear autocorrelation function of the stock market return. The model achieves this result without any dependence between consecutive price changes. It proves the crucial role of inter-event times in volatility clustering phenomenon.
I. INTRODUCTION

In recent years we observe a rapid increase of interest in point processes and their applications, especially in financial data modeling [1]. This kind of stochastic process considers events occurring irregularly in time and describes times between these events and their dependencies. Two of the most popular ones are Autoregressive Conditional Duration (ACD) [2] and Hawkes model [3, 4]. The canonical versions of both models include short-range dependencies (for ACD see [2, 5–10], for Hawkes see [11–19]). However, both of them were extended to describe long-range memory (for ACD see [20–30], for Hawkes see [31–41]).

In many cases, not only we observe some events occurring irregularly in time, but also some value that can be measured in these discrete moments. The high-frequency transaction data from the stock market is an excellent example. We observe events - transactions occurring in specific moments, but also with each transaction, we can relate quantities of price and volume. Of course, in such cases, the inter-event times can be modeled as a point process.

The first formalism to describe dynamics of observed value in discrete times was continuous-time random walk (CTRW) [42, 43] proposed in 1965 by Montroll and Weiss. The CTRW models have found many applications, including astrophysics, geophysics, econophysics, and sociophysics. For a more detailed review, see [44]. In the canonical CTRW, both increments of the observed process and waiting times (inter-event times) are i.i.d. random variables. The example trajectory of such a process is shown in Fig. [1]. All kinds of random walks, starting with normal diffusion, by anomalous diffusion (both subdiffusion and superdiffusion), to Levy flights, can be described within the CTRW formalism. It can be obtained by using specific distributions of waiting times or increments (especially with heavy tails) and by considering memory in waiting times, increments or dependence between them. The CTRW
models with correlated increments were initially proposed to describe lattice gases \cite{45,47}. However, recently they have been used to model financial high-frequency data \cite{48,60}. On the other hand, the CTRW models with correlated waiting times are not well-studied. Except for a few papers \cite{52,61,62}, these models were not considered nor used to model empirical data. That seems to be surprising, in the light of the latest popularity of point processes such as ACD or Hawkes process. The aim of this work is to some extent fill this gap, and to propose a new CTRW formalism, which is capable of considering dependencies in inter-event times. The main point is to model long-range memories in waiting times. As shown later, the formalism proposed in this paper, despite its simplicity, is general enough to fill the gap and can be compared with empirical data and other models.

The paper is organized as follows. In section II we present the motivation from the financial data, particularly by looking at the non-stationary and stationarized data. Next, we propose a way to include dependencies between the waiting times III, especially the long-range memory. Then, we solve the CTRW model with correlated waiting times IV by calculating its moments and the autocorrelation function (ACF) of increments. We also fit our model to tick-by-tick transaction data from the Warsaw Stock Exchange in section V. Finally, we sum up our work in section VI. Additionally, we include Appendix A to clarify the mathematical methods to obtain our results.

II. MEMORIES IN FINANCIAL DATA

As mentioned in the introduction, the model presented in this paper is directly motivated by high frequency, tick-by-tick data from the stock market \cite{?}. However, as dependencies as described here can be observed in different areas, its possible applications exceed this narrow motivation.

At first, let us remind two basic stylized facts observed in the majority of stock
FIG. 1. The example trajectory of the continuous-time random walk (CTRW) consisting of jumps of the process values $\Delta x_n$ preceded by the waiting times $\Delta t_n$. In the canonical CTRW $\Delta t_n$ and $\Delta x_n$ are i.i.d. random variables drawn from the distributions $\psi (\Delta t_n)$ and $h (\Delta x_n)$ respectively. In this paper, we are considering the CTRW model with long term dependence in the series of consecutive waiting times $\Delta t_1, \Delta t_2, \ldots, \Delta t_n$. Both of these facts are related to time ACF. In the case of time ACF of log returns, we observe short-term negative autocorrelation. In the other case of time ACF of absolute values of log returns, we observe slowly decaying positive autocorrelation. The latter is considered a reminiscence of the volatility clustering phenomenon [63]. The CTRW models were already used to describe high-frequency stock market data. Taking into account the so-called bid-ask bounce phenomenon allows us to reproduce the first stylized fact of short-term negative autocorrelation of log returns [58, 64, 65]. In this type of the models $\Delta t_n$ were still i.i.d. variables, but $\Delta x_n$ were depended on $\Delta x_{n-1}$. Unfortunately, models considering only this type of dependencies turned out to be unable to describe time ACF of price changes absolute values [60]. Technically, it is possible to obtain the CTRW model reproducing both stylized facts, but it requires power-law waiting-time distribution $\psi (\Delta t)$. However, this solution cannot be satisfying as we can obtain waiting-time distribution directly from the empirical data of inter-transaction times. It turns out that this distribution is
not even close to a power-law one [58]. The source of the second stylized fact is not in the distributions \( h(\Delta x), \psi(\Delta t) \), but in the dependence between consecutive \( \Delta x \) and \( \Delta t \). As we were interested in the construction of the CTRW model accurately describing the second stylized fact, we asked ourselves which dependencies we have to take it into account and which ones can be neglected.

Let us start with simple analysis of step ACF of series \( \Delta t_n \) and \( |\Delta x_n| \). We observe almost power-law memories in waiting times and price changes absolute values, see Fig. 2a. Firstly (lag \( \lesssim 3 \) ) autocorrelation of \( |\Delta x_n| \) is higher, but for lag \( > 3 \) autocorrelation of \( \Delta t_n \) is more significant. This result suggests that in the limit of long times, the dependence between waiting times may be more critical than dependence between price changes. To verify this hypothesis, we performed a simple shuffling test. We compared the time ACF of price changes absolute values for four samples of time series. First of them is the original time series of tick-by-tick transaction data. We expected to observe the second stylized fact in this case. The second time series keeps the price changes \( \Delta x_n \) in the original order but shuffles the waiting times \( \Delta t_n \). This way we obtained the time series keeping all dependencies between price changes \( \Delta x_n \) but removing dependencies between waiting times \( \Delta t_n \). In the third time series, we kept the original waiting times \( \Delta t_n \) but shuffles the price changes \( \Delta x_n \). In the last, fourth time series, both \( \Delta t_n \) and \( \Delta x_n \) were shuffled. Let us emphasise that all four time series have the same, unchanged distributions \( \psi(\Delta t_n) \) and \( h(\Delta x_n) \). The results are shown in Fig. 2b. As expected, we do observe slow, almost power-law decay of time ACF for the first, empirical time series. Surprisingly, removing dependencies between waiting times does not change the time ACF in the limit of \( t \to 0 \) but significantly increases the slope of the decay of time ACF in the long-term. On the other hand, removing dependencies between price changes decreases the time ACF by almost constant factor but does not change the slope of the decay. Removing of all dependencies still exhibits positive time ACF, which is the result of the
non-exponential empirical distribution of waiting times.

These observations motivate us to construct CTRW model with long-range dependencies between waiting times, which should be able to reproduce slowly decaying ACF as in financial data. It even suggests that it is impossible to model empirical data without using long-term dependencies in waiting times.

The results shown above implicate that we should focus on long-term dependencies within waiting times to properly model stock price behavior. In the previous section, we analyzed step ACF for lags up to 100 and time ACF for times up to 1000 s. Such limits were chosen due to the length of the trading sessions (around 8 hours or 1000 trades). Unfortunately, these limits are not long enough to detect power-law dependencies. The only way to increase these limits is by joining all sessions into one
sequence. In this procedure, we merged the end of one session with the beginning of the following one (we omit overnight price changes). We have to deal with the fact that these two periods of the sessions are not quite the same. Moreover, we observe intraday non-stationarity in financial data. The session begins with relatively short inter-transaction times and a relatively high standard deviation of price changes. Usually, up to the middle of the session, average inter-trade times increase, and the standard deviation of price changes decreases. The situation reverts close to the end of the session. This phenomenon is called the lunch effect. We use the canonical method to remove intraday non-stationarity is by dividing each waiting time by the corresponding average waiting time, depending on the time from the trading beginning and the day of a week. The comparison of step ACF of waiting times for non-stationarized and stationarized data is presented in Fig. 3a. As a result of this procedure, we obtain the power-law decay over four orders of magnitude of lag. In Fig. 3b, we present the time ACF of price changes absolute values for stationarized data, which also exhibit power-law decay over four orders of magnitude of time. It seems reasonable to study the relationship between decay exponents of these autocorrelations. Fortunately, the model studied in this paper gives a strict answer to this question.

III. PROCESS OF TIMES

Let us now focus on the series of inter-transaction times $\Delta t_1, \Delta t_2, \ldots, \Delta t_n, \ldots$. We are looking now for the point process to describe this series, which will be suitable to be used later within CTRW of prices. For this reason, we prefer analytically solvable models. Moreover, we would like to keep the original distribution of inter-transaction times $\psi(\Delta t_n)$ and observe the power-law step ACF, as shown in Fig. 3a. Even these two simple conditions exclude ACD models and Hawkes processes from
FIG. 3. All intraday data (waiting times and corresponding price changes) were joined into one data set. a) The plot shows normalized step ACF of $\Delta t$ for non-stationary and stationarized case. Stationarizing procedure is described in the main text. b) The plot of normalized time ACF of $|\Delta x|$ with stationarized waiting times. Both stationarized autocorrelations decay like a power-law with similar slope.

our considerations. We are not interested in ACD models, as the power-law ACF can be obtained only within the fractional extension. In the Hawkes process, both waiting time distribution and autocorrelation depend on the memory kernel. Therefore they cannot be set independently. By setting the memory kernel, which reproduces the empirical waiting time distribution $\psi(\Delta t_n)$, we obtain specific step ACF, without any degree of freedom to change it. This feature of Hawkes processes significantly hampers its use in the description of empirical data.

The solution to our search for a suitable point process turned out to be surprisingly simple. We noticed that we are already dealing with the stochastic process satisfying our requirements. It is called the continuous-time random walk (CTRW). Within the canonical CTRW, values of the process are represented by a spatial variable, and the time is continuous. Adapting the CTRW to the role of a point process requires the value of the process to represent waiting time and the subordinated time to be discrete. As in the CTRW values of the process are constant during the waiting times in the case of the discrete-time the analog of waiting time can be considered
as the number of repetitions \( \nu_i \) of the same value. The example trajectory of such adapted, subordinated CTRW is shown in Fig. 4.

We require the waiting times \( \Delta t_n \) (values of the process in the discrete subordinated time \( n \)) to come from distribution \( \psi(\Delta t_n) \ (\Delta t_n > 0) \), with finite mean \( \langle \Delta t \rangle \). Let the \( \nu_i \) be the number of repeats of the same waiting times (drawn independently for each series of repetitions). Let \( \nu_i \) be the i.i.d. random variables with distribution \( \omega(\nu) \). In general, it can be any distribution, but to recreate power-law step ACF of waiting times we will focus on fat-tailed distribution with finite first moment \( \langle \nu \rangle \).

For example, zeta distribution with parameter \( \rho \):

\[
\omega_\rho(n) = n^{-\rho}/\zeta(\rho); \quad \zeta(\rho) = \sum_{n=1}^{\infty} n^{-\rho}, \rho > 1, \tag{1}
\]

where \( \zeta(\rho) \) is Riemann’s zeta function. Its mean \( \langle \omega \rangle = \frac{\zeta(\rho-1)}{\zeta(\rho)} \) for \( \rho > 2 \) and the variance is finite for \( \rho > 3 \). The CDF is given by \( H_{n,\rho}/\zeta(\rho) \), where \( H_{n,\rho} = \sum_{k=1}^{n} k^{-\rho} \) is generalized harmonic number. Let us introduce \( \Omega(\nu) \) as a sojourn probability, so \( \Omega(\nu) = 1 - \frac{H_{n-1,\rho}}{\zeta(\rho)} \) for zeta distribution.

The typical way to describe a stochastic process is to compute its soft propagator \( P(\Delta t; n|\Delta t_0, 0) \). It is defined as the conditional probability density that the process value, which was initially (at \( n = 0 \)) in the origin value (\( \Delta t = \Delta t_0 \)), at time \( n \) is equal to \( \Delta t \). Soft propagator can be expressed intuitively:

\[
P(\Delta t; n|\Delta t_0, 0) = \delta(\Delta t - \Delta t_0)\Omega_1(n) + [1 - \Omega_1(n)]\psi(\Delta t). \tag{2}
\]

The first term is the probability, that the process value will stay constant (in \( \Delta t_0 \)) after \( n \) jumps. The second term indicates that there will be process value jump with probability \( 1 - \Omega_1(n) \), so new process values will be completely independent and it will come from the distribution \( \psi(\Delta t) \). The \( \Omega_1(n) \) is sojourn probability from
The example trajectory of the subordinatory CTRW, which value correspond to the waiting times $\Delta t_n$ used in the main CTRW process. Process values are $\Delta t_1, \Delta t_2, \ldots, \Delta t_n, \ldots$, which are repeated respectively $\nu_1, \nu_2, \ldots, \nu_n, \ldots$ times. Number of repeats $\nu_i$ are drown from the distribution $\omega(\nu_i)$.

stationarized (case of the first jump) $\omega_1(n)$:

\[
\omega_1(n) = \frac{\sum_{n'=1}^{n} \omega(n + n')}{\sum_{n''=0}^{\infty} \sum_{n'=1}^{n} \omega(n'' + n')} = \frac{\sum_{n'=1}^{n} \omega(n + n')}{\sum_{n=1}^{\infty} n\omega(n)} = \frac{\sum_{n'=n+1}^{\infty} \omega(n')}{\langle \omega \rangle}
\]

\[
\Omega_1(n) = \frac{\sum_{i=n}^{\infty} \sum_{n'=i+1}^{\infty} \omega(n')}{\langle \omega \rangle} = \frac{\sum_{i=1}^{n} i\omega(i + n)}{\langle \omega \rangle} = \frac{\langle \omega \rangle - n\Omega(n + 1) - \sum_{i=1}^{n} i\omega(i)}{\langle \omega \rangle}
\]

(3)

Restricting ourselves to $\omega(n)$ in the form of zeta distribution we can obtain

\[
\Omega_1(n) = 1 - \frac{n}{\langle \omega' \rangle} + \frac{nH_{n,\rho}}{\zeta(\rho - 1)} - \frac{H_{n,\rho-1}}{\zeta(\rho - 1)},
\]

(4)

propagator and step autocorrelation of waiting times $\Delta t$. The step autocovariance of times $\Delta t$ can be expressed as

\[
cov(n) = \langle \Delta t_i \Delta t_{i+n} \rangle - \langle \Delta t_i \rangle \langle \Delta t_{i+n} \rangle = \langle \Delta t_i \Delta t_{i+n} \rangle - \langle \Delta t \rangle^2,
\]

(5)
where symbol \( \langle \ldots \rangle \) means taking the average. Please note that \( \Delta t_{i+n} = \Delta t_i \) with probability \( p = \Omega_1(n) \). With probability \( 1 - p \), the \( \Delta t_i \) is independent. This lead to

\[
cov(n) = p \langle \Delta t^2 \rangle + (1 - p) \langle \Delta t \rangle^2 = \sigma^2_{\Delta t} p = \sigma^2_{\Delta t} \Omega_1(n). \tag{6}
\]

We are interested in the asymptotic for of autocorrelation for \( n \gg 1 \), We can use to following approximation (Theorem 12.21 from [69])

\[
\zeta(\rho) - H_{n,\rho} \approx \frac{n^{1-\rho}}{\rho - 1}. \tag{7}
\]

Finally, we obtain normalized step ACF

\[
corr(n) = \frac{cov(n)}{cov(0)} \approx \frac{n^{-(\rho-2)}}{\zeta(\rho-1)(\rho-2)(\rho-1)}. \tag{8}
\]

As we required the step ACF of waiting times decays like a power-law and the decay exponent is \( -(\rho - 2) \). It is worth emphasizing that, even considering only \( \rho > 2 \), required for the existence of a finite average number of repetitions, we can obtain any value of the decay exponent.

IV. THE PRIMARY PROCESS. THE CTRW IN CTRW

Now we are ready to define the primary process, in our financial application describing, for instance, the price of the financial asset in time. This process is also the CTRW with two key properties

- changes of the process value \( \Delta x \) are i.i.d. random variables from the distribution \( h(\Delta x) \), with finite variance \( \sigma_x^2 \) (and thus finite first two moments \( \mu_1 \) and \( \mu_2 \)).
- Waiting times \( t \) come from subordinated CTRW process of times described in
the previous section[III]

This way we obtain in some sense the CTRW in CTRW. Of course, this name is not a strict one, as the subordinated time is discrete, so usage of continuous in the context of the subordinated process is a small misuse. Please note, that we do not assume any dependence within the series of consecutive changes of the process value $\Delta x_1, \Delta x_2, \ldots \Delta x_n$. We do not make any further assumptions about the shape of the distributions $h(\Delta x)$. Using the symmetric distribution, with vanishing mean, we will approximately describe the process of price changes in time. This symmetric distribution will allow us to obtain i.e., the autocorrelation of the price changes. Following [60], if as $h(\Delta x)$ we use only the positive half of the distribution multiplied by 2, we deal with the case of non-zero drift ($\mu_1 \neq 0$) and obtain artificial, monotonically increasing process. The autocorrelation of increment of this process will represent the autocorrelation of price changes absolute values, which is particularly interesting in our analysis.

We managed to obtain the soft propagator and the following characteristics of the proper CTRW process. Although the mathematical methods of finding the propagator are exciting, they are not crucial for understanding the main result of this manuscript. The details of the calculations can be found in Appendix [A]. Here we present the selected results, namely two first moments and time autocorrelation of changes, in the limit of long times ($t \to \infty$). We consider analytical terms $(t, t^2, t^3, \ldots)$ and the most significant power-law term when $\rho$ is non-integer.

The first moment of the process for $t \to \infty$ can be approximated as

$$m_1(t) \approx \frac{\mu_1}{\langle \Delta t \rangle} t + \mu_1 \frac{\alpha \{\psi\}}{\Gamma(4 - \rho)} t^{3-\rho}, \quad \rho \in (2; 4),$$

where $\Gamma(\cdot)$ is gamma function. For $\rho > 4$, there is still power-law term dependent on $\mu_1 t^{3-\rho}$, but its amplitude is not known. The most important term is typical linear
behavior, but we observe additional power-law term. The second moment can be written in the form

$$m_2(t) \approx \mu_1^2 \left( \frac{t}{\langle \Delta t \rangle} \right)^2 + \sigma_x^2 \frac{t}{\langle \Delta t \rangle} + \mu_1^2 \beta \{ \psi \} \frac{t}{\langle \Delta t \rangle} + \mu_1^2 \frac{\gamma \{ \psi \}}{\Gamma(5 - \rho)} t^{4 - \rho}, \quad \rho \in (2; 5). \quad (10)$$

From the first two moments of the process, we calculated the process variance (still considering only analytical and the most important power-law term)

$$\sigma^2(t) \approx \left( \sigma_x^2 + \mu_1^2 \beta \{ \psi \} \right) \frac{t}{\langle \Delta t \rangle} + \mu_1^2 \frac{\gamma \{ \psi \}}{\Gamma(5 - \rho)} t^{4 - \rho}, \quad \rho \in (2; 5). \quad (11)$$

It is worth to mention, that for variance the power-law term from the second moment is more important than power-law term from the first moment. We can observe normal diffusion for $\rho > 3$. However, there is superdiffusion in the case of $\rho \in (2; 3)$. We obtain ballistic diffusion in the limit $\rho \to 2$.

Having the first two moments, one can calculate ACF of changes for stationary process

$$C(t) = \frac{1}{2} \frac{\partial^2 m_2(t)}{\partial t^2} - \left( \frac{\partial m_1(t)}{\partial t} \right)^2 \Rightarrow C(t) \approx \mu_1^2 \kappa \{ \psi, \rho \} t^{2 - \rho}, \quad (12)$$

where $\kappa \{ \psi, \rho \} = \frac{1}{\Gamma(3 - \rho)} \left( \frac{\gamma \{ \psi \}}{2} - \frac{2 \alpha \{ \psi \}}{\langle \Delta t \rangle} \right)$, for $\rho \in (2; 4)$. Interestingly, for $\mu_1 \neq 0$, we observe the power-law behavior of time ACF of changes in the limit of long times.

Now we would like to use the proposed process to describe high-frequency financial data. It turns out that this model can describe both stylized facts concerning autocorrelation functions. If we use symmetric distributions $h(x)$ as the distribution of returns is symmetric, we obtain $\mu_1 = 0$ and quickly decaying autocorrelations of returns. We can also calculate the nonlinear autocorrelations function of the modules of returns, by calculating linear autocorrelation functions of the other process that uses only positive half of the distribution $h(x)$. This way, we obtain slow, power-law
decay of the autocorrelation. As we assumed only one type of memory in this process, introduced unrealistically by the distribution $\omega(\nu)$, it is not surprising that the ACF of the process depends on it. The surprising is the fact that in this model, the exponent of the decay of the nonlinear time ACF is the same as in the step ACF of waiting times. This fact motivated us to compare these two values for empirical financial data we use.

V. EMPIRICAL RESULTS

In the proposed model, the decay exponents (slopes on the log-log plot) of step ACF of $\Delta t$ and time ACF of $|\Delta x|$ are the same. We can compare this prediction with empirical data. Of course, in empirical data we also observe long-term positive step ACF of $|\Delta x|$, which was not included in our model, so finally we can expect that the slope of time ACF of $|\Delta x|$ should be slightly higher than the slope of step ACF of $\Delta t$. As long-term non-linear autocorrelation is usually interpreted as a reminiscence of the volatility clustering phenomena, it is interesting to check what part of the observed volatility clustering effect can be explained only by memory between inter-trade times. We present results for 5 most traded stocks from the Warsaw Stock Exchange below (ordered by number of transactions) with average inter-trade time not greater than 30 seconds.

We see that our model does not fit perfectly with empirical data, but it is enough to estimate the slope of time ACF with accuracy around 10%.

VI. CONCLUSIONS

We introduced the new Continuous Time Random Walks (CTRW) model with the long-term memory within consecutive waiting times. We assume that the consec-
TABLE I. Table with fitted slopes of empirical stationarized step ACF of waiting times and time ACF of price changes absolute values for 5 most liquid stocks from WSE. The time ACF slopes are close to corresponding step ACF slopes. The analysis was performed on the tick-by-tick market data from the public domain database [?]. The data covers the period 2013-01-03 till 2017-07-14. For instance, the data set for KGHM contains 3 096 625 transactions.

| Company | Step ACF $\Delta t$ slope | Time ACF $|\Delta x|$ slope |
|---------|----------------|----------------|
| KGHM   | $-0.25 \pm 0.04$ | $-0.25 \pm 0.02$ |
| PKOBP  | $-0.33 \pm 0.08$ | $-0.30 \pm 0.02$ |
| PZU    | $-0.26 \pm 0.03$ | $-0.28 \pm 0.04$ |
| PGE    | $-0.33 \pm 0.07$ | $-0.36 \pm 0.03$ |
| PEKAO  | $-0.33 \pm 0.04$ | $-0.37 \pm 0.04$ |

...
Appendix A:

In Appendix, we sketch the solution for calculating the moments of the process in the limit of long times. All increments of the process $\Delta x$ are independent, so firstly we will focus only on the number of jumps. We calculate the probability $P_n(t)$ for $n \geq 0$, which is the probability that it will be exactly $n$ jumps up to time $t$. $P_0(t)$ can be obtained directly from the definition, as the probability of no jumps in the time $t$ is

$$P_0(t) = \Psi(t) \Rightarrow \tilde{P}_0(s) = \tilde{\Psi}(s), \quad (A1)$$

where $\Psi(t)$ is the sojourn probability for the waiting time distribution. For $n \geq 1$ the process will be described by the number of series of waiting times $k$, waiting times in each series $t_i$ and the number of repetitions of waiting times in each series $\nu_i$. Particularly, the equations $\nu_1 + \nu_2 + \cdots + \nu_k = n$ and $\nu_1 t_1 + \nu_2 t_2 + \cdots + \nu_k t_k \leq t$ must hold. The soft propagator $P_n(t)$ for $n \geq 1$ can be written as the sum of two parts:

1. the $k$-th series of waiting times $t_k$ repeated $\nu_k$ times ended before time $t$ and the process is still in the same position (the next waiting time will be from the new series),

2. the process is during the series of WT $t_k$, which was repeated $\nu_k$ times so far, in the time $t$. 

16
For simplicity of notation, let’s redefine $\Omega(\nu) = \sum_{n=\nu+1}^{\infty} \omega(n)$.

\[
P_n(t) = \sum_{k=1}^{n} \sum_{\nu_1,\ldots,\nu_k} \int_{0<\delta t} \psi(t_1) \ldots \psi(t_k) \Psi(\delta t) \omega(\nu_1) \ldots \omega(\nu_k) dt_1 \ldots dt_k
+ \sum_{k=1}^{n} \sum_{\nu_1,\ldots,\nu_k} \int_{0<\delta t} \psi(t_1) \ldots \psi(t_k) \omega(\nu_1) \ldots \omega(\nu_{k-1}) \Omega(\nu_k) dt_1 \ldots dt_k,
\]

(A2)

where $\delta t = t - \sum t_i \nu_i$. Next, we calculate the Laplace transform ($t \to s$) and Z transform ($n \to z$) to obtain

\[
\tilde{P}(z; s) = \tilde{P}_0(s) + \tilde{P}_{z}(s) = \frac{1}{s} \frac{1}{s - f(z; s)} \left[ 1 + \tilde{F}(z; s) - z(\tilde{F}(z; s) + \tilde{f}(z; s)) \right],
\]

(A3)

where $\tilde{f}(z; s) = \sum_{\nu=1}^{\infty} z^{-\nu} \tilde{\psi}(s\nu) \omega(\nu)$ and analogically $\tilde{F}(z; s) = \sum_{\nu=1}^{\infty} z^{-\nu} \tilde{\psi}(s\nu) \Omega(\nu)$. Let’s notice that the full soft propagator with included jumps can be easily expressed as the Z transform of $\tilde{P}_n$ at the point $z = \tilde{h}(k)^{-1}$

\[
\tilde{P}(k; s) = \sum_{n=0}^{\infty} \tilde{P}_n \tilde{h}^n(k) = \tilde{P}(z; s) \bigg|_{z=\tilde{h}(k)^{-1}} = \frac{1 + \tilde{F}(\tilde{h}(k)^{-1}; s) - \tilde{h}(k)^{-1}(\tilde{F}(\tilde{h}(k)^{-1}; s) + \tilde{f}(\tilde{h}(k)^{-1}; s))}{s - \tilde{F}(\tilde{h}(k)^{-1}; s)}.
\]

(A4)

The first two moments of the process can be calculated as the derivatives of the
propagator at the point \( k = 0 \):

\[
\tilde{m}_1(s) = -i \frac{\partial \tilde{P}(k; s)}{\partial k} \bigg|_{k=0} = \frac{\mu_1 J_0 + j_0}{s(1 - j_0)},
\]

\[
\tilde{m}_2(s) = -\frac{\partial^2 \tilde{P}(k; s)}{\partial k^2} \bigg|_{k=0} = \frac{2\mu_1^2 j_1(J_0 + j_0) + (1 - j_0)(J_1 + j_1 - J_0 - j_0)}{(1 - j_0)^2}
\]

\[+ \frac{\mu_2 J_0 + j_0}{s(1 - j_0)}, \tag{A5}\]

where we introduced

\[
\hat{j}_n = j(n; s) = \sum_{\nu=1}^{\infty} \nu^n \tilde{\psi}(s\nu) \omega(\nu), \quad J_n = J(n; s) = \sum_{\nu=1}^{\infty} \nu^n \tilde{\psi}(s\nu) \Omega(\nu). \tag{A6}\]

Next, we focus on the specific power-law memory. We set the distribution of the number of repeats to be Zipf’s distribution with the parameter \( \rho \): \( \omega(\nu) = \frac{\nu^{-\rho}}{\zeta(\rho)}, \rho > 2 \).

The parameter \( \rho \) has to be bigger than two because the distribution of the number of the repeats must have finite mean not to break ergodicity. Also, we expand the moments into the series assuming very small \( s \) (so for the long times). To do that we need the expansions of the \( \hat{j}(n; s) \) and \( J(n; s) \) for the \( n = \{0, 1\} < (\rho - 1) \). One can express \( \hat{j}(n; s) \) as the power-law sum:

\[
\hat{j}(n; s) = \frac{1}{\zeta(\rho)} \sum_{\nu=1}^{\infty} \tilde{\psi}(s\nu) \nu^{-(\rho-n)} = s^{\rho-n-1} \zeta(\rho) I_{(s\nu)} \tag{A7}\]

The behaviour of the sum \( I \) can be estimated by the integrals

\[
\int_{2s}^{\infty} \tilde{\psi}(x)x^{-(\rho-n)}dx < I < \int_{s}^{\infty} \tilde{\psi}(x)x^{-(\rho-n)}dx. \tag{A8}\]
Therefore, we can approximate the sum $I$ into series and finally obtain

$$j(n; s) = C_n s^{\rho-n-1} + C_n^0 s + C_n^1 s^2 + C_n^2 s^3 + \cdots. \quad (A9)$$

One can calculate

$$C_n^0 = j(n; 0) = \frac{1}{\zeta(\rho)} \sum_{\nu=1}^{\infty} \nu^{-(\rho-n)} = \frac{\zeta(\rho-n)}{\zeta(\rho)} \geq 1. \quad (A10)$$

Moreover we can notice that

$$\frac{C_1^0}{C_0^0} = -\frac{1}{\langle \Delta t \rangle}. \quad (A11)$$

Similarly we approximated

$$J(n; s) = D_n s^{\rho-n-2} + D_n^0 s + D_n^1 s^2 + D_n^2 s^3 + \cdots. \quad (A12)$$

The constant term are:

$$D_0^0 = \frac{\zeta(\rho-1)}{\zeta(\rho)} - 1 = C_1^0 - C_0^0, \quad D_1^0 = \frac{\zeta(\rho-2) - \zeta(\rho-1)}{2\zeta(\rho)} = \frac{C_2^0 - C_1^0}{2}. \quad (A13)$$

This gives us the form of the first moment

$$\tilde{m}_1(s) \approx \frac{\mu_1}{s} \left( C_1^0 + D_0 s^{\rho-2} \right) \frac{C_0^0 s^{\rho-2} - 1}{s C_0^1} = -\frac{\mu_1}{s^2} \frac{C_1^0}{C_0^1} - \frac{\mu_1}{s^{4-\rho}} \frac{D_0 + \frac{C_0 C_1^0}{C_0^1}}{C_0^1}. \quad (A14)$$

concerning only terms increasing with time ($s^{-\alpha}, \alpha > 1$): analytical and the most important power-law one. Switching to the time variables, we obtain:

$$m_1(t) = \mathcal{L}^{-1} [\tilde{m}_1(s)] \approx \frac{\mu_1}{\langle \Delta t \rangle} t - \frac{\mu_1}{C_0^1 \Gamma(4-\rho)} t^{3-\rho}. \quad (A15)$$

19
The second moment can be expressed as
\[
\tilde{m}_2(s) \approx \frac{2\mu_1^2}{\langle \Delta t \rangle^2} s^{-3} - \mu_1^2 \frac{4C_0^2 + 3C_0^1 \langle \Delta t \rangle + 2C_0^1 \langle \Delta t \rangle + 2D_0^1 \langle \Delta t \rangle + C_2^0 \langle \Delta t \rangle^2}{2C_0^1 \langle \Delta t \rangle^2} s^{-2}
\]
- \mu_1^2 \frac{D_0 + C_1 - D_1 + 2C_0}{C_0^1 \langle \Delta t \rangle} s^{\theta - 5} + \frac{\mu_2}{\langle \Delta t \rangle} s^{-2}.
\]
(A16)

This gives us the variance in the time domain, presented in the main text.

[1] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events: for Insurance and Finance*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013.

[2] R. F. Engle and J. R. Russell. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66:1127–1162, 1998.

[3] Hawkes A.G. Spectra of some self-exciting and mutually-exciting point processes. *Biometrika*, 58:83, 1971.

[4] Hawkes A.G. Point spectra of some mutually-exciting point processes. *J. R. Stat. Soc. B*, 33:438, 1971.

[5] Alfonso Dufour and Robert F. Engle. Time and the price impact of a trade. *The Journal of Finance*, 55(6):2467–2498, 2000.

[6] Robert F Engle and Joe Lange. Predicting vnet: A model of the dynamics of market depth. *Journal of Financial Markets*, 4(2):113 – 142, 2001.

[7] Robert F. Engle and Jeffrey R. Russell. Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model. *Journal of Empirical Finance*, 4(2):187 – 212, 1997. High Frequency Data in Finance, Part 1.
[8] Luc Bauwens and Pierre Giot. The logarithmic acd model: An application to the bid-ask quote process of three NYSE stocks. *Annales d'Économie et de Statistique*, 60:117–149, 2000.

[9] Joachim Grammig and Kai-Oliver Maurer. Non-monotonic hazard functions and the autoregressive conditional duration model. *The Econometrics Journal*, 3(1):16–38, 2000.

[10] Maria Pacurar. Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Surveys*, 22(4):711–751, 2008.

[11] Adamopoulos L. Hawkes A. G. Cluster models for earthquakes—regional comparisons. *Bull. Int. Stat. Inst.*, 45:454, 1973.

[12] Ogata Y. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Ann. Inst. Stat. Math.*, 30:243, 1978.

[13] Brillinger D.R. The identification of point process systems. *Ann. Probab.*, 3:909, 1975.

[14] Oakes D. The markovian self-exciting process. *J. Appl. Probab.*, 12:69, 1975.

[15] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.

[16] Bowsher C.G. Modelling security market events in continuous time: Intensity based, multivariate point process models. *J. Econometrics*, 141:876, 2007.

[17] McNeil A.J. Chavez-Demoulin V., Davison A.C. Estimating value-at-risk: A point process approach. *Quant. Finance*, 5:227, 2005.

[18] Patrick Hewlett. Clustering of order arrivals, price impact and trade path optimisation. In *In Workshop on Financial Modeling with Jump processes, Ecole Polytechnique*, 2006.

[19] Large J. Measuring the resiliency of an electronic limit order book. *J. Financ. Markets,*
[20] Jan Beran, Yuanhua Feng, and Sucharita Ghosh. Modelling long-range dependence and trends in duration series: an approach based on efarima and esemifar models. *Statistical Papers*, 56(2):431–451, May 2015.

[21] Joanna Jasiak. Persistence in intertrade durations. *Finance*, 19:166–195, 1999.

[22] Menelaos Karanasos. The statistical properties of exponential acd models. *Quantitative and Qualitative Analysis in Social Sciences*, 2:29–49, 2008.

[23] Jan Beran and Yuanhua Feng. Iterative plug-in algorithms for semifar models: Definition, convergence, and asymptotic properties. *Journal of Computational and Graphical Statistics*, 11(3):690–713, 2002.

[24] Jan Beran and Yuanhua Feng. Semifar models—a semiparametric approach to modelling trends, long-range dependence and nonstationarity. *Computational Statistics & Data Analysis*, 40(2):393 – 419, 2002.

[25] Rohit Deo, Mengchen Hsieh, and Clifford M. Hurvich. Long memory in intertrade durations, counts and realized volatility of nyse stocks. *Journal of Statistical Planning and Inference*, 140(12):3715 – 3733, 2010. Special Issue in Honor of Emanuel Parzen on the Occasion of his 80th Birthday and Retirement from the Department of Statistics, Texas A&M University.

[26] Rohit Deo, Clifford M. Hurvich, Philippe Soulier, and Yi Wang. Conditions for the propagation of memory parameter from durations to counts and realized volatility. *Econometric Theory*, 25(3):764–792, 2009.

[27] Wei Sun, Svetlozar Rachev, Frank J. Fabozzi, and Petko S. Kalev. Fractals in trade duration: capturing long-range dependence and heavy tailedness in modeling trade duration. *Ann. Finance*, 4(2):217–241, 2008.

[28] Eric Ghysels and Joanna Jasiak. Garch for irregularly spaced financial data: The acd-garch model. *Studies in Nonlinear Dynamics & Econometrics*, 2.4, 1998.
[29] Nikolaus Hautsch. *Econometrics of financial high-frequency data*. Berlin: Springer, 2012.

[30] Saranjeet Kaur Bhogal and Ramanathan Thekke Variyam. Conditional duration models for high-frequency data: A review on recent developments. *Journal of Economic Surveys*, 33(1):252–273, 2019.

[31] Alan G. Hawkes. Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2):193–198, 2018.

[32] Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 01(01):1550005, 2015.

[33] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. II*. Probability and its Applications (New York). Springer, New York, second edition, 2008. General theory and structure.

[34] McGill J.A. Chavez-Demoulin V. High-frequency financial data modeling using hawkes processes. *J. Banking Finance*, 36:3415, 2012.

[35] E. Bacry, K. Dayri, and J. F. Muzy. Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data. *The European Physical Journal B*, 85(5):157, May 2012.

[36] Emmanuel Bacry and Jean-François Muzy. Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, 14(7):1147–1166, 2014.

[37] Vladimir Filimonov and Didier Sornette. Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Phys. Rev. E*, 85:056108, May 2012.

[38] V. Filimonov and D. Sornette. Apparent criticality and calibration issues in the hawkes self-excited point process model: application to high-frequency financial data. *Quantitative Finance*, 15(8):1293–1314, 2015.

[39] Stephen J. Hardiman, Nicolas Bercot, and Jean-Philippe Bouchaud. Critical reflexivity in financial markets: a hawkes process analysis. *The European Physical Journal
B, 86(10):442, Oct 2013.

[40] Stephen J. Hardiman and Jean-Philippe Bouchaud. Branching-ratio approximation for the self-exciting hawkes process. *Phys. Rev. E*, 90:062807, Dec 2014.

[41] Thibault Jaisson and Mathieu Rosenbaum. Limit theorems for nearly unstable hawkes processes. *Ann. Appl. Probab.*, 25(2):600–631, 04 2015.

[42] Point processes extended to fit this phenomenon are called marked point processes [15].

[43] E. W. Montroll and G. H. Weiss. Random walks on lattices. II. *Journal of Mathematical Physics*, 6(2):167–181, 1965.

[44] Ryszard Kutner and Jaume Masoliver. The continuous time random walk, still trendy: fifty-year history, state of art and outlook. *The European Physical Journal B*, 90(3):50, Mar 2017.

[45] R. Kutner. Correlated hopping in honeycomb lattice: tracer diffusion coefficient at arbitrary lattice gas concentration. *Journal of Physics C: Solid State Physics*, 18(34):6323, 1985.

[46] K.W. Kehr, R. Kutner, and K. Binder. Diffusion in concentrated lattice gases. self-diffusion of noninteracting particles in three-dimensional lattices. *Physical Review B*, 23(10):4931–4945, 1981.

[47] J. W. Haus and K. W. Kehr. Random walk model with correlated jumps: Self-correlation function and frequency-dependent diffusion coefficient. *Journal of Physics and Chemistry of Solids*, 40(12):1019 – 1025, 1979.

[48] E. Scalas, R. Gorenflo, and F. Mainardi. Fractional calculus and continuous-time finance. *Physica A: Statistical Mechanics and its Applications*, 284(1-4):376 – 384, 2000.

[49] F. Mainardi, M. Raberto, R. Gorenflo, and E. Scalas. Fractional calculus and continuous-time finance ii: the waiting-time distribution. *Physica A: Statistical Me-
M. Raberto, E. Scalas, and F. Mainardi. Waiting-times and returns in high-frequency financial data: an empirical study. *Physica A: Statistical Mechanics and its Applications*, 314(1-4):749–755, 2002.

E. Scalas, R. Gorenflo, and F. Mainardi. Uncoupled continuous-time random walks: Solution and limiting behavior of the master equation. *Physical Review E*, 69(1):011107, 2004.

E. Scalas. The application of continuous-time random walks in finance and economics. *Physica A: Statistical Mechanics and its Applications*, 362(2):225–239, 2006.

R. Kutner and F. Świtała. Stochastic simulations of time series within Weierstrass-Mandelbrot walks. *Quantitative Finance*, 3(3):201–211, 2003.

J. Masoliver, M. Montero, and G. H. Weiss. Continuous-time random-walk model for financial distributions. *Physical Review E*, 67(2):021112, 2003.

P. Repetowicz and P. Richmond. Modeling share price evolution as a continuous time random walk (ctrw) with non-independent price changes and waiting times. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):108–111, 2004. Applications of Physics in Financial Analysis 4 (APFA4).

J. Masoliver, M. Montero, J. Perelló, and G. H. Weiss. The continuous time random walk formalism in financial markets. *Journal of Economic Behavior & Organization*, 61(4):577–598, 2006.

J. Masoliver, M. Montero, J. Perello, and G. H. Weiss. The CTRW in finance: Direct and inverse problems with some generalizations and extensions. *Physica A: Statistical Mechanics and its Applications*, 379(1):151–167, 2007.

Tomasz Gubiec and Ryszard Kutner. Backward jump continuous-time random walk: An application to market trading. *Phys. Rev. E*, 82:046119, Oct 2010.

Tomasz Gubiec and Ryszard Kutner. Continuous-time random walk with multi-step
memory: an application to market dynamics. *The European Physical Journal B*, 90(11):228, Nov 2017.

[60] Jarosław Klamut and Tomasz Gubiec. Directed continuous-time random walk with memory. *The European Physical Journal B*, 92(4):69, Apr 2019.

[61] Vincent Tejedor and Ralf Metzler. Anomalous diffusion in correlated continuous time random walks. *Journal of Physics A: Mathematical and Theoretical*, 43(8):082002, 2010.

[62] Aleksei V. Chechkin, Michael Hofmann, and Igor M. Sokolov. Continuous-time random walk with correlated waiting times. *Phys. Rev. E*, 80:031112, Sep 2009.

[63] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.

[64] M. Montero and J. Masoliver. Nonindependent continuous-time random walks. *Physical Review E*, 76(6):061115, 2007.

[65] T. Gubiec and R. Kutner. Share price evolution as stationary, dependent continuous-time random walk. *Acta Physica Polonica, A*, 117(4):669 – 672, 2010.

[66] R. Gençay, M. Dacorogna, U.A. Müllner, O. Pictet, and R. Olsen. *An Introduction to High-Frequency Finance*. Elsevier Science, 2001.

[67] Michael M. Dacorogna, Ulrich A. Müllner, Robert J. Nagler, Richard B. Olsen, and Olivier V. Pictet. A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance*, 12(4):413 – 438, 1993.

[68] Ruey S. Tsay. *Analysis of financial time series*. Wiley series in probability and statistics. Wiley-Interscience, Hoboken, NJ, 2. ed. edition, 2005.

[69] Tom M. Apostol. *Introduction to Analytic Number Theory*. Springer-Verlag, New York, 1976. Undergraduate Texts in Mathematics.

[70] Richard T. Baillie, Tim Bollerslev, and Hans Ole Mikkelsen. Fractionally integrated
generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30, 1996.

[71] Dmitri Koulikov. Modeling sequences of long memory non-negative covariance stationary random variables. *William Davidson Institute Working Paper No. 493*, 2003.