

A non-linear double stochastic model of return in financial markets

Vygintas Gontis, Julius Ruseckas and Aleksejus Kononovičius
Institute of Theoretical Physics and Astronomy of Vilnius University
Lithuania

1. Introduction

Volatility clustering, evaluated through slowly decaying auto-correlations, Hurst effect or $1/f$ noise for absolute returns, is a characteristic property of most financial assets return time series Willinger et al. (1999). Statistical analysis alone is not able to provide a definite answer for the presence or absence of long-range dependence phenomenon in stock returns or volatility, unless economic mechanisms are proposed to understand the origin of such phenomenon Cont (2005); Willinger et al. (1999). Whether results of statistical analysis correspond to long-range dependence is a difficult question and subject to an ongoing statistical debate Cont (2005).

Extensive empirical analysis of the financial market data, supporting the idea that the long-range volatility correlations arise from trading activity, provides valuable background for further development of the long-ranged memory stochastic models Gabaix et al. (2003); Plerou et al. (2001). The power-law behavior of the auto-regressive conditional duration process Sato (2004) based on the random multiplicative process and its special case the self-modulation process Takayasu (2003), exhibiting $1/f$ fluctuations, supported the idea of stochastic modeling with a power-law probability density function (PDF) and long-range memory. Thus the agent based economic models Kirman & Teyssiere (2002); Lux & Marchesi (2000) as well as the stochastic models Borland (2004); Gontis et al. (2008; 2010); Queiros (2007) exhibiting long-range dependence phenomenon in volatility or trading volume are of great interest and remain an active topic of research.

Properties of stochastic multiplicative point processes have been investigated analytically and numerically and the formula for the power spectrum has been derived Gontis & Kaulakys (2004). In the more recent papers Kaulakys et al. (2006); Kaulakys & Alaburda (2009); Ruseckas & Kaulakys (2010) the general form of the multiplicative stochastic differential equation (SDE) was derived in agreement with the model earlier proposed in Gontis & Kaulakys (2004). Since Gontis & Kaulakys (2004) a model of trading activity, based on a SDE driven Poisson-like process, was presented Gontis et al. (2008) and in the most recent paper Gontis et al. (2010) we proposed a double stochastic model, whose return time series yield two power-law statistics, i.e., the PDF and the power spectral density (PSD) of absolute return, mimicking the empirical data for the one-minute trading return in the NYSE.

In this chapter we present theoretical arguments and empirical evidence for the non-linear double stochastic model of return in financial markets. With empirical data from NYSE and Vilnius Stock Exchange (VSE) demonstrating universal scaling of return statistical properties,

which is also present in the double stochastic model of return Gontis et al. (2010). The sections in this chapter follow the chronology of our research papers devoted to the stochastic modeling of financial markets. In the second sections we introduce multiplicative stochastic point process reproducing $1/f^\beta$ noise and discuss its possible application as the stochastic model of financial market. In the section 3 we derive multiplicative SDE statistically equivalent to the introduced point process. Further, in the section 4 we propose a Poisson-like process driven by multiplicative SDE. More sophisticated version of SDE reproducing statistics of trading activity in financial markets is presented in the section 5 and empirical analysis of high frequency trading data from NYSE in the section 6. Section 7 introduces the stochastic model with a q -Gaussian PDF and power spectrum $S(f) \sim 1/f^\beta$ and the section 8 the double stochastic model of return in financial market. We present scaled empirical analysis of return in New York and Vilnius stock exchanges in comparison with proposed model in the sections 9. Short conclusions of the most recent research results is presented in the section 10.

2. $1/f$ noise: from physics to financial markets

The PSD of a large variety of different evolutionary systems at low frequencies have $1/f$ behavior. $1/f$ noise is observed in condensed matter, river discharge, DNA base sequence structure, cellular automata, traffic flow, economics, financial markets and other complex systems with the evolutionary elements of self-organization (see, e.g., a bibliographic list of papers by Li (2009)). Considerable amount of such systems have fractal nature and thus their statistics exhibit scaling. It is possible to define a stochastic model system exhibiting fractal statistics and $1/f$ noise, as well. Such model system may represent the limiting behavior of the dynamical or deterministic complex systems, explaining the evolution of the complexity into chaotic regime.

Let us introduce a multiplicative stochastic model for the time interval between events in time series, defining in such a way the multiplicative point process. This model exhibits the first order and the second order power-law statistics and serves as the theoretical description of the empirical trading activity in the financial markets Gontis & Kaulakys (2004).

First of all we consider a signal $I(t)$ as a sequence of the random correlated pulses

$$I(t) = \sum_k a_k \delta(t - t_k) \quad (1)$$

where a_k is a contribution to the signal of one pulse at the time moment t_k , e.g., a contribution of one transaction to the financial data. Signal (1) represents a point process used in a large variety of systems with the flow of point objects or subsequent actions. When $a_k = \bar{a}$ is constant, the point process is completely described by the set of times of the events $\{t_k\}$ or equivalently by the set of inter-event intervals $\{\tau_k = t_{k+1} - t_k\}$.

Various stochastic models of τ_k can be introduced to define a such stochastic point process. In the papers Kaulakys & Meškauskas (1998); Kaulakys (1999; 2000) it was shown analytically that the relatively slow Brownian fluctuations of the inter-event time τ_k yield $1/f$ fluctuations of the signal (1). In the generalized version of the model Gontis & Kaulakys (2004) we have introduced a stochastic multiplicative process for the inter-event time τ_k ,

$$\tau_{k+1} = \tau_k + \gamma \tau_k^{2\mu-1} + \tau_k^\mu \sigma \varepsilon_k. \quad (2)$$

Here the inter-event time τ_k fluctuates due to the external random perturbation by a sequence of uncorrelated normally distributed random variable $\{\varepsilon_k\}$ with zero expectation and unit

variance, σ denotes the standard deviation of the white noise and $\gamma \ll 1$ is a damping constant. Note that from the big variety of possible stochastic processes we have chosen the multiplicative one as it yields multifractal intermittency and power-law PDF. Certainly, in Eq. (2) the τ_k diffusion has to be restricted in some area $0 < \tau_{\min} < \tau_k < \tau_{\max}$. Multiplicativity is specified by μ (pure multiplicativity corresponds to $\mu = 1$, while other values of might be considered as well).

The iterative relation (2) can be rewritten as Langevin SDE in k -space, inter-event space,

$$d\tau_k = \gamma\tau_k^{2\mu-1} + \sigma\tau_k^\mu dW_k. \tag{3}$$

Here we interpret k as continuous variable while W_k defines the Wiener noise in inter-event space.

Steady state solution of the stationary Fokker-Planck equation with zero flow, corresponding to (3), gives the probability density function for τ_k in the k -space (see, e.g., Gardiner (1986))

$$P_k(\tau_k) = C\tau_k^\alpha = \frac{\alpha + 1}{\tau_{\max}^{(\alpha+1)} - \tau_{\min}^{(\alpha+1)}} \tau_k^\alpha, \quad \alpha = 2\gamma/\sigma^2 - 2\mu. \tag{4}$$

The steady state solution (4) assumes Ito convention involved in the relation between expressions (2), (3) and (4) and the restriction for the diffusion $0 < \tau_{\min} < \tau_k < \tau_{\max}$. In the limit $\tau_{\min} \rightarrow 0$ and $\tau_{\max} \rightarrow \infty$ the explicit expression of the signal's $I(t)$ PSD $S_\mu(f)$ was derived in Gontis & Kaulakys (2004):

$$S_\mu(f) = \frac{C\bar{\alpha}^2}{\sqrt{\pi\bar{\tau}}(3 - 2\mu)f} \left(\frac{\gamma}{\pi f} \right)^{\frac{\alpha}{3-2\mu}} \frac{\Gamma(\frac{1}{2} + \frac{\alpha}{3-2\mu})}{\cos(\frac{\pi\alpha}{2(3-2\mu)})}. \tag{5}$$

Equation (5) reveals that the multiplicative point process (2) results in the PSD $S(f) \sim 1/f^\beta$ with the scaling exponent

$$\beta = 1 + \frac{2\gamma/\sigma^2 - 2\mu}{3 - 2\mu}. \tag{6}$$

Analytical results (5) and (6) were confirmed with the numerical calculations of the PSD according to equations (1) and (2).

Let us assume that $a \equiv 1$ and the signal $I(t)$ counts the transactions in financial markets. In that case the number of transactions in the selected time window τ_d , defined as $N(t) = \int_t^{t+\tau_d} I(t)dt$, measures the trading activity. PDF of N for the pure multiplicative model, with $\mu = 1$, can be expressed as, for derivation see Gontis & Kaulakys (2004),

$$P(N) = \frac{C'\tau_d^{2+\alpha}(1 + \gamma N)}{N^{3+\alpha}(1 + \frac{\gamma}{2}N)^{3+\alpha}} \sim \begin{cases} \frac{1}{N^{3+\alpha}}, & N \ll \gamma^{-1}, \\ \frac{1}{N^{5+2\alpha}}, & N \gg \gamma^{-1}. \end{cases} \tag{7}$$

Numerical calculations confirms the obtained analytical result (7).

In the case of pure multiplicativity, $\mu = 1$, the model has only one parameter, $2\gamma/\sigma^2$, which defines scaling of the PSD, the power-law distributions of inter-event time and the number of deals N per time window. The model proposed with the adjusted parameter $2\gamma/\sigma^2$ nicely describes the empirical PSD and the exponent of power-law long range distribution of the trading activity N in the financial markets, see Gontis & Kaulakys (2004) for details.

Ability of the model to simulate $1/f$ noise as well as to reproduce long-range power-law statistics of trading activity in financial markets promises wide interpretation and application of the model. Nevertheless, there is an evident need to introduce Poisson-like flow of trades in high frequency time scales of financial markets.

3. Power-law statistics arising from the nonlinear stochastic differential equations

In the previous section we introduced the stochastic multiplicative point process, which was proposed in Gontis & Kaulakys (2004), presented a formula for the PSD and discussed a possible application of the model to reproduce the long-range statistical properties of financial markets. The same long-range statistical properties pertaining to the more general ensemble of stochastic systems can be derived from the SDE or by the related Fokker-Plank equation. Supposing that previously introduced multiplicative point process reflects long-range statistical properties of financial markets, we feel the need to derive multiplicative SDE statistically equivalent to the introduced point process. It would be very nice if the SDE was applicable towards the modeling of financial markets as well.

Transition from the occurrence number, k , space SDE to the actual time, t , space in the SDE(3) can be done according to the relation $dt = \tau_k dk$. This transition yields

$$d\tau = \gamma\tau^{2\mu-2} + \sigma\tau^{\mu-1/2}dW. \quad (8)$$

One can transform variables in the SDE (8) from inter-event time, τ , to the average intensity of the signal, $I(t)$, which itself can be expressed as $x = a/\tau$, or to the number of events per unit time interval $n = 1/\tau$. Applying Ito transform of variables to the SDE (8) gives new SDE for x

$$dx = (\sigma^2 - \gamma)\frac{x^{4-2\mu}}{a^{3-2\mu}} + \frac{\sigma x^{5/2-\mu}}{a^{3/2-\mu}}dW. \quad (9)$$

One can introduce scaled time

$$t_s = \frac{\sigma^2}{a^{3-2\mu}}t, \quad (10)$$

and some new parameters

$$\eta = \frac{5}{2} - \mu, \quad \lambda = 2\left(\frac{\gamma}{\sigma^2} + \frac{3}{2} - \mu\right), \quad (11)$$

in order to obtain the class of Ito SDE

$$dx = \left(\eta - \frac{\lambda}{2}\right)x^{2\eta-1} + x^\eta dW_s. \quad (12)$$

Eq. (12), as far as it corresponds to the point process discussed in the previous section, should generate the power-law distributions of the signal intensity,

$$P(x) \sim x^{-\lambda}, \quad (13)$$

and $1/f^\beta$ noise,

$$S(f) \sim \frac{1}{f^\beta}, \quad \beta = 1 - \frac{3 - \lambda}{2\eta - 2}. \quad (14)$$

In some cases time series obtained from SDE (12) may diverge, thus hampering numerical calculations. In the real systems some diffusion restriction mechanisms are present, thus restricting diffusion of SDE solutions seems rather natural. One can introduce the exponential restriction into SDE (12) setting distribution densities Gardiner (1986):

$$P(x) \sim x^{-\lambda} \exp \left\{ - \left(\frac{x_{min}}{x} \right)^m - \left(\frac{x}{x_{max}} \right)^m \right\}. \tag{15}$$

In that case SDE (12) is rewritten as

$$dx = \left(\eta - \frac{\lambda}{2} + \frac{m}{2} \left\{ \left(\frac{x_{min}}{x} \right)^m - \left(\frac{x}{x_{max}} \right)^m \right\} \right) x^{2\eta-1} + x^\eta dW_s, \tag{16}$$

where m is parameter responsible for sharpness of restriction.

Many numerical simulations were performed to prove validity of power-law statistics (14)-(15) for the class of SDE (16) Kaulakys et al. (2006). Recently (see Ruseckas & Kaulakys (2010)) it was shown that power-law statistics (14)-(15) can be derived directly from the SDE, without relying on the formalization of point processes (namely model discussed in previous section). This, more general, derivation serves as additional justification of equations and provides further insights into the origin of power-law statistics.

4. Fractal point process driven by the nonlinear stochastic differential equation

In the previous section starting from the point process (1) we derived the class of nonlinear SDE (12) or, with limits towards diffusion, (16) . One can consider the appropriate SDE as an initial model of long-range power-law statistics driving another point process in microscopic level. In Gontis & Kaulakys (2006; 2007) we proposed to model trading activity in financial markets as Poisson-like process driven by nonlinear SDE.

From the SDE class (16) one can draw SDE for the number of point events or trades per unit time interval, n , which would be expressed as

$$dn = \left\{ \eta - \frac{\lambda}{2} + \frac{m}{2} \left(\frac{n_0}{n} \right)^m \right\} n^{2\eta-1} dt_s + n^\eta dW_s. \tag{17}$$

The Poisson-like flow of events can be introduced by conditional probability of inter-event time, τ ,

$$\varphi(\tau|n) = n \exp(-n\tau). \tag{18}$$

Note that here τ is measured in scaled time, t_s , units and the expectation of instantaneous inter-event time, for instantaneous n , is

$$\langle \tau \rangle_n = \int_0^\infty \tau \varphi(\tau|n) d\tau = \frac{1}{n}. \tag{19}$$

The long-range PDF of n , time series obtained from the with Eq. (17) related Fokker-Plank equation, has an explicit form:

$$P_m^{(t)}(n) = \frac{m}{n_0 \Gamma(\frac{\lambda-1}{m})} \left(\frac{n_0}{n} \right)^\lambda \exp \left(- \left(\frac{n_0}{n} \right)^m \right). \tag{20}$$

Similarly the long-range PDF of τ can be written as

$$P_m^{(t)}(\tau) = \tau \int_0^\infty n \varphi(\tau|n) P(n) dn = \frac{m}{\tau_0 \Gamma(\frac{\lambda-1}{m})} \frac{\tau}{\tau_0} \int_0^\infty x^{2-\lambda} \exp\left(-\frac{1}{x^m} - x \frac{\tau}{\tau_0}\right) dx, \quad (21)$$

here we have performed substitution of parameters $\tau_0 = \frac{1}{n_0}$. The integral in (21) has an explicit form, when $m = 1$

$$P_{m=1}(\tau) = \frac{2}{\tau_0 \Gamma(\lambda-1)} \left(\frac{\tau}{\tau_0}\right)^{\frac{\lambda-1}{2}} K_{\lambda-3}\left(2\sqrt{\frac{\tau}{\tau_0}}\right) \quad (22)$$

Here $K_n(z)$ is the modified Bessel function of the second kind. When $\tau \rightarrow \infty$, we get

$$P_{m=1}(\tau) \approx \frac{\sqrt{\pi}}{\tau_0 \Gamma(\lambda-1)} \left(\frac{\tau}{\tau_0}\right)^{\frac{\lambda-3}{2}} \exp\left(-2\sqrt{\frac{\tau}{\tau_0}}\right). \quad (23)$$

The integral in (21) can be expressed via special functions, when $m = 2$. However, we can obtain asymptotic behavior for small and large $\frac{\tau}{\tau_0}$. When $\frac{\tau}{\tau_0} \rightarrow \infty$, using the method of the steepest descent we get

$$P_{m=2}^{(t)}(\tau) = \frac{2^{\frac{10-\lambda}{3}}}{\tau_0 \Gamma(\frac{\lambda-1}{2})} \sqrt{\frac{\pi}{3}} \left(\frac{\tau}{\tau_0}\right)^{\frac{\lambda-1}{3}} \exp\left(-3\left(\frac{\tau}{2\tau_0}\right)^{\frac{2}{3}}\right), \quad (24)$$

while in case of $\frac{\tau}{\tau_0} \rightarrow 0$ one can obtain

$$P_{m=2}^{(t)}(\tau) \rightarrow \begin{cases} \frac{2\Gamma(3-\lambda)}{\tau_0 \Gamma(\frac{\lambda-1}{2})} \left(\frac{\tau}{\tau_0}\right)^{\lambda-2} & 1 < \lambda < 3 \\ \frac{\Gamma(\frac{\lambda-3}{2})}{\tau_0 \Gamma(\frac{\lambda-1}{2})} \frac{\tau}{\tau_0} & \lambda > 3 \end{cases} \quad (25)$$

5. Fractal trading activity of financial market driven by the nonlinear stochastic differential equation

We will investigate how previously introduced modulated Poisson stochastic point process can be adjusted to the empirical trading activity, defined as number of transactions in the selected time window τ_d . In order to obtain the number of events, N , in the selected time window, τ_d , one has to integrate the stochastic signal Eq. (17) in the corresponding time interval. We denote the integrated number of events, N , as

$$N(t, \tau_d) = \int_t^{t+\tau_d} n(t') dt' \quad (26)$$

and call it the trading activity in case of the financial markets.

Detrended fluctuation analysis Plerou et al. (2000) is one of the ways to analyze the second order statistics related to the autocorrelation of trading activity. The exponents of the detrended fluctuation analysis, ν , obtained by fits for each of the 1000 US stocks show a relatively narrow spread of ν around the mean value $\nu = 0.85 \pm 0.01$ Plerou et al. (2000). We use relation $\beta = 2\nu - 1$ between the exponents ν of the detrended fluctuation analysis and the exponents β of the PSD Beran (1994) and in this way define the empirical value of the exponent for

the power spectral density $\beta = 0.7$. Our analysis of the Vilnius stock exchange (VSE) data confirmed that the PSD of trading activity is the same for various liquid stocks even for the emerging markets Gontis & Kaulakys (2004). The histogram of exponents obtained by fits to the cumulative distributions of trading activities of 1000 US stocks Plerou et al. (2000) gives the value of exponent $\lambda = 4.4 \pm 0.05$ describing the power-law behavior of the trading activity. Empirical values of $\beta = 0.7$ and $\lambda = 4.4$ confirm that the time series of the trading activity in real markets are fractal with the power law statistics. Time series generated by stochastic process (17) are fractal in the same sense.

Nevertheless, we face serious complications trying to adjust model parameters to the empirical data of the financial markets. For the pure multiplicative model, setting $\mu = 1$ or $\eta = 3/2$, we have to take $\lambda = 2.7$ to get $\beta = 0.7$, while empirical λ value being noticeably different - 4.4, i.e. it is impossible to reproduce the empirical PDF and PSD with the same exponent of multiplicativity η . We have proposed possible solution of this problem in our publications Gontis & Kaulakys (2004) deriving PDF for the trading activity N , see Eq. (7). When $N \gg \gamma^{-1}$ one can obtain exactly the required values of $\lambda = 5 + 2\alpha = 4.4$ and $\beta = 0.7$ for $\gamma_\sigma = \frac{\gamma}{\sigma^2} = 0.85$. Despite model being able to mimic empirical data under certain conditions, we cannot accept it as the sufficiently accurate model of the trading activity since the empirical power law distribution is achieved only for very high values of the trading activity. This discrepancy provides insight to the mechanism of the power law distribution converging to the normal distribution through increasing values of the exponent, though empirically observed power law distribution in wide area of N values cannot be reproduced. Let us notice here that the desirable power law distribution of the trading activity with the exponent $\lambda = 4.4$ may be generated by the model (17) with $\eta = 5/2$. Moreover, only the smallest values of τ or high values of n contribute to the power spectral density of trading activity Kaulakys et al. (2006). Thus we feel incentive to combine the stochastic processes with two values of μ or η : (i) $\mu \simeq 0$ or $\eta \simeq 5/2$ for the main area of τ and n diffusion and (ii) $\mu = 1$ or $\eta \simeq 3/2$ for the lowest values of τ or highest values of n . Therefore, we introduce a new SDE for n , which includes two powers of the multiplicative noise,

$$dn = \left[\left(\frac{5}{2} - \frac{\lambda}{2} \right) + \frac{m}{2} \left(\frac{n_0}{n} \right)^m \right] \frac{n^4}{(n\epsilon + 1)^2} dt + \frac{n^{5/2}}{(n\epsilon + 1)} dW, \tag{27}$$

where a new parameter ϵ defines crossover between two areas of n diffusion. The corresponding iterative equation for τ_k in such a case is expressed as

$$\tau_{k+1} = \tau_k + \left[\left(\frac{\lambda + 2}{2} - \eta \right) - \frac{m}{2} \left(\frac{\tau}{\tau_0} \right)^m \right] \frac{\tau_k}{(\epsilon + \tau_k)^2} + \frac{\tau_k}{\epsilon + \tau_k} \epsilon_k, \tag{28}$$

where ϵ_k denotes uncorrelated normally distributed random variable with the zero expectation and unit variance.

Eqs. (27) and (28) define related stochastic variables n and τ , respectively, and they should reproduce the long-range statistical properties of the trading activity and of waiting time in the financial markets. We verify this by the numerical calculations. In Figure 1 we present the PSD calculated for the equivalent processes (a)-(27) and (b)-(28) (see Gontis & Kaulakys (2004) for details of calculations). This approach reveals the structure of the PSD in wide range of frequencies and shows that the model exhibits not one, but two rather different power laws with the exponents $\beta_1 = 0.34$ and $\beta_2 = 0.74$. In Figure 1 we also present the distributions of trading activity (c) and (d), which now have correct exponents. From many numerical calculations performed with the multiplicative point processes we can conclude that combination

of two power laws of spectral density arise only when the multiplicative noise is a crossover of two power laws as in Eqs. (27) and (28).

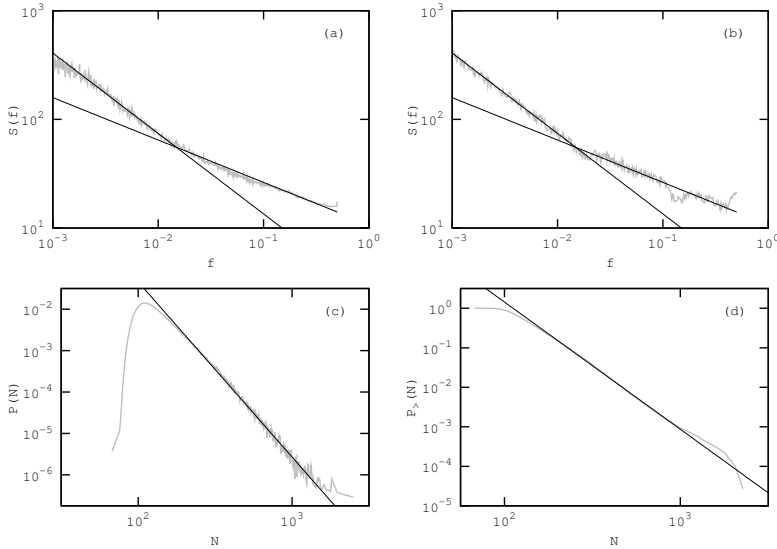


Fig. 1. (a) PSD, $S(f)$, of the τ dependent flow generated by Eq. (28), (b) $S(f)$ calculated from n time series generated by Eq. (27), (c) PDF, $P(N)$, of number of trades, obtained by integrating signal, within time window of $\tau_d = 100s$, N , generated by Eq. (28), (d) corresponding inverse cumulative distribution, $P_>(N)$, of N . Statistical properties are represented by gray curves, while black lines approximate their power laws: (a) and (b) black lines give approximation $S(f) \sim 1/f^{\beta_{1,2}}$ with $\beta_1 = 0.34$ and $\beta_2 = 0.74$, (c) black line gives approximation $P(N) \sim N^{-\lambda}$ with exponent $\lambda = 4.3$, (d) black line gives approximation $P_>(N) \sim N^{-\lambda}$ with exponent $\lambda = 3.3$. Statistical properties were obtained using corresponding model parameter values: $\lambda = 4.3$; $\epsilon = 0.07$; $\tau_0 = 1$; $m = 6$.

Thus as of now we have introduced the complete set of equations defining the stochastic model of the trading activity in the financial markets. We have proposed this model following our growing interest in the stochastic fractal point processes Gontis & Kaulakys (2004); Kaulakys et al. (2005; 2006). Our objective to reproduce, in details, statistics of trading activity is the cause for rather complicated form of the SDE (27) and thus there is low expectation of analytical results. Therefore we focus on the numerical analysis and direct comparison of the model with the empirical data.

In order to achieve general description of statistics for different stocks we introduce the scaling into Eq. (27) utilizing scaled rate $x = n/n_0$ and $\epsilon' = \epsilon n_0$. After the substitution Eq. (27) becomes

$$dx = \left[\left(\frac{5}{2} - \frac{\lambda}{2} \right) + \frac{m}{2} x^{-m} \right] \frac{x^4}{(x\epsilon' + 1)^2} dt + \frac{x^{5/2}}{(x\epsilon' + 1)} dW. \quad (29)$$

We have eliminated parameter n_0 as it is specific for each stock. By doing so we also decrease number of model parameters to three, which must be defined from the empirical data of trading activity in the financial markets.

One can solve Eq. (29) using the method of discretization. Thus we introduce variable step of integration $\Delta t = h_k = \kappa^2/x_k$, and the differential equation (29) transforms into the set of difference equations

$$x_{k+1} = x_k + \kappa^2 \left[\left(\frac{5}{2} - \frac{\lambda}{2} \right) + \frac{m}{2} x_k^{-m} \right] \frac{x_k^3}{(x_k e' + 1)^2} + \kappa \frac{x_k^2}{(x_k e' + 1)} \varepsilon_k, \tag{30}$$

$$t_{k+1} = t_k + \kappa^2/x_k \tag{31}$$

with $\kappa \ll 1$ being small parameter and ε_k defining Gaussian noise with zero mean and unit variance .

With the substitution of variables, namely $\tau = 1/n$, one can transform Eq. (27) into

$$d\tau = \left[\frac{\lambda - 3}{2} - \frac{m}{2} \left(\frac{\tau}{\tau_0} \right)^m \right] \frac{1}{(\varepsilon + \tau)^2} dt + \frac{\sqrt{\tau}}{\varepsilon + \tau} dW \tag{32}$$

with limiting time $\tau_0 = 1/n_0$. Further we argue that this form of driving SDE is more suitable for the numerical analysis. First of all, the powers of variables in this equation are lower, but the main advantage is that the Poissonian-like process can be included into the procedure of numerical solution of SDE. As we did with SDE for n we should also introduce a scaling of Eq. (32). It is done by defining the non-dimensional scaled time $t_s = t/\tau_0$, scaled inter-trade time $y = \tau/\tau_0$ and $e' = \varepsilon/\tau_0$. After those transformations Eq. (32) becomes

$$dy = \frac{1}{\tau_0^2} \left[\frac{\lambda - 3}{2} - \frac{m}{2} y^m \right] \frac{1}{(e' + y)^2} dt_s + \frac{1}{\tau_0} \frac{\sqrt{y}}{e' + y} dW_s. \tag{33}$$

As in the real discrete market trading we can choose the instantaneous inter-trade time y_k as a step of numerical calculations, $h_k = y_k$, or even more precisely as the random variables with the exponential distribution $P(h_k) = 1/y_k \exp(-h_k/y_k)$. We obtain iterative equation resembling tick by tick trades in the financial markets,

$$y_{k+1} = y_k + \frac{1}{\tau_0^2} \left[\frac{\lambda - 3}{2} - \frac{m}{2} y_k^m \right] \frac{h_k}{(e' + y_k)^2} + \frac{1}{\tau_0} \frac{\sqrt{y_k h_k}}{e' + y_k} \varepsilon_k. \tag{34}$$

In this numerical procedure the sequence of $1/y_k$ gives the modulating rate, n , and the sequence of h_k is the Poissonian-like inter-trade times. Seeking higher precision one can use the Milshtein approximation for Eq. (33) instead of Eq. (34).

6. Analysis of empirical stock trading data

Previously, see Gontis et al. (2008), we have analyzed the tick by tick trades of 26 stocks on NYSE traded for 27 months from January, 2005. In this chapter we will briefly discuss main results and valuable conclusions, providing important insights, of the empirical analysis presented in Gontis et al. (2008). Empirical analysis is very important as starting from it we can adjust the parameters of the Poisson-like process driven by SDE Eq. (27) or Eq. (34) to numerically reproduce the empirical trading statistics.

An example of the empirical histogram of τ_k and $N(t, \tau_d)$ and the PSD of IBM trade sequence are shown on Figure 2. The histograms and PSD of the sequences of trades for all 26 stocks are similar to IBM shown on Fig. 2. From the histogram, $P(\tau_k)$, we can obtained model parameter τ_0 value for every stock. One can define the exponent λ' from the power-law tail $P(N) \sim N^{-\lambda'}$

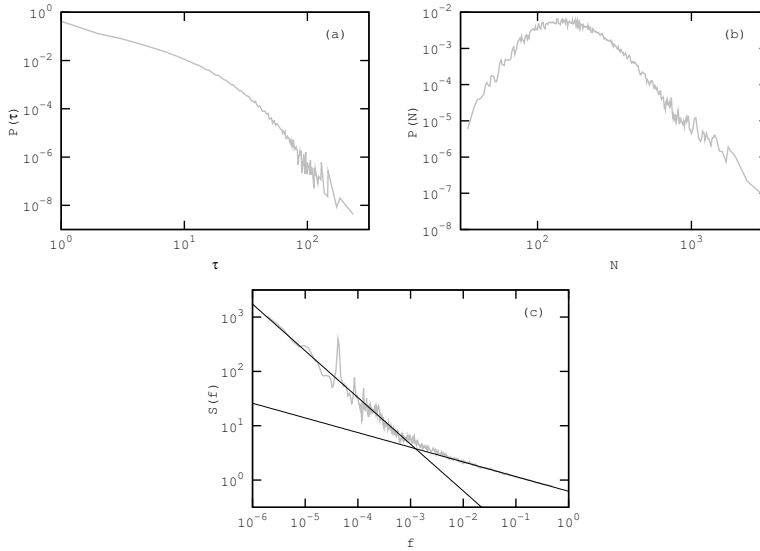


Fig. 2. Trading statistics of IBM stocks. (a) Empirical histogram of the inter-trade time τ_k sequence, $P(\tau)$; (b) histogram of trading activity, $P(N)$, calculated in the time interval $\tau_d = 600$ s; (c) PSD, $S(f)$, of the sequence of trades (gray curve), straight black lines approximate PSD $S(f) \sim 1/f^{\beta_{1,2}}$ with $\beta_1 = 0.33$ and $\beta_2 = 0.94$.

of N histogram. The PSD exhibits two scaling exponents β_1 and β_2 if approximated by power-law $S(f) \sim f^{-\beta_{1,2}}$.

Empirical values of β_1 and β_2 fluctuate around 0.3 and 0.9, respectively, same behavior is observed in different stochastic model realizations. The crossover frequency f_c of two power-laws exhibits some fluctuations around the value $f_c \approx 10^{-3}$ Hz as well. One can observe considerable fluctuations of the exponent λ' around the mean value 4.4. We would like to note that the value of histogram exponent, λ' , for integrated trading activity N is higher than for n , as λ' increases with higher values of time scale τ_d .

From the point of view of the proposed model parameter τ_0 is specific for every stock and reflects the average trading intensity in the calm periods of stock exchange. In previous section we have shown that one can eliminate these specific differences in the model by scaling transform of Eq. (32) arriving to the nondimensional SDE (33) and its iterative form (34). These equations and parameters $\sigma' = \sigma/\tau_0$, λ , ϵ' and $m = 2$ define model, which has to reproduce, in details, power-law statistics of the trading activity in the financial markets. From the analysis based on the research of fractal stochastic point processes Gontis & Kaulakys (2004; 2006; 2007); Kaulakys et al. (2005; 2006) and by fitting the numerical calculations to the empirical data we arrive at the conclusion that model parameters should be set as $\sigma' = 0.006$, $\lambda = 4.3$, $\epsilon' = 0.05$ in order to achieve best results. In Figure 3 we have presented statistical properties obtained from our model using aforementioned parameter values - PDF of the sequence of $\tau_k = h_k$, (a), and the PSD of the sequence of trades as point events, (b).

For every selected stock one can easily scale the model sequence of inter-trade times $\tau_k = h_k$ by empirically defined τ_0 to get the model sequence of trades for this stock. One can scale the

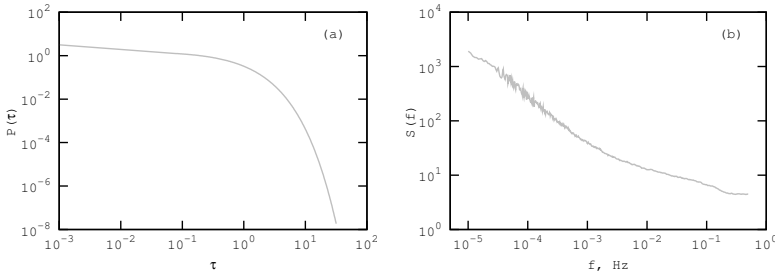


Fig. 3. Distribution of the Poisson-like inter-event times $\tau_k = h_k$, (a), and power spectrum, (b), of the sequence of point events calculated from Eq. (34) with the adjusted parameters $m = 2$, $\sigma' = 0.006$, $\lambda = 4.3$, $\epsilon' = 0.05$, $\tau_0 = 1$.

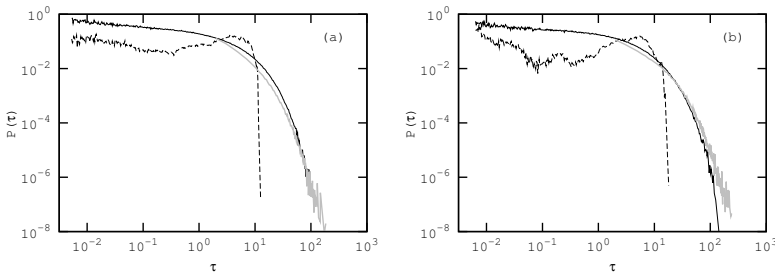


Fig. 4. Distribution of inter-trade times, τ , for (a) IBM and (b) MMM stocks; empirical histogram, gray curve, modeled Poisson-like distribution, black solid curve, distribution of driving $\tau = y_k$ in Eq. (34), black dashed curve. Model parameters are the same as in Fig. 3. $\tau_0 = 5$ s for IBM and $\tau_0 = 7.25$ s for MMM stocks.

model power spectrum $S(f)$ by $1/\tau_0^2$ for getting the model power spectrum $S_{\text{stock}}(f)$ for the selected stock $S_{\text{stock}}(f) = S(f)/\tau_0^2$.

Previously we have proposed the iterative Eq. (34) as quite accurate stochastic model of trading activity in the financial markets. Nevertheless, one has to admit that real trading activity often has considerable trend as number of shares traded and the whole activity of the markets increases. This might have considerable influence on the empirical long-range distributions and power spectrum of the stocks in consideration. The trend has to be eliminated from the empirical data for the detailed comparison with the model. Only few stocks from the selected list have stable trading activity in the considered period.

In Figure 4, Figure 5 and Figure 6 we compare the model statistical properties with the empirical statistics of the stocks with stable trading activity. As we show in Figure 4, the model Poisson-like distribution can be easily adjusted to the empirical histogram of inter-trade time, with $\tau_0 = 5$ s for IBM trade sequence and with $\tau_0 = 7.25$ s for MMM trading. The comparison with the empirical data is limited by the available accuracy, 1 s, of stock trading time t_k . The probability distribution of driving $\tau = y_k$ Eq. (34), dashed line, illustrates different market behavior in the periods of the low and high trading activity. The Poissonian nature of the stochastic point process hides these differences by considerable smoothing of the PDF. In Figure 5 one can see that the long-range memory properties of the trading activity reflected in

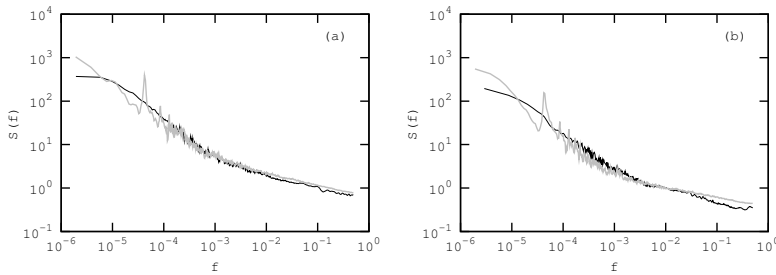


Fig. 5. Modeled, black curves, and empirical, gray curves, PSD of trading activity, N , for (a) IBM and (b) MMM stocks. Parameters are the same as in Fig. 3. $\tau_0 = 5\text{s}$ for IBM and $\tau_0 = 7.25\text{s}$ for MMM stocks.

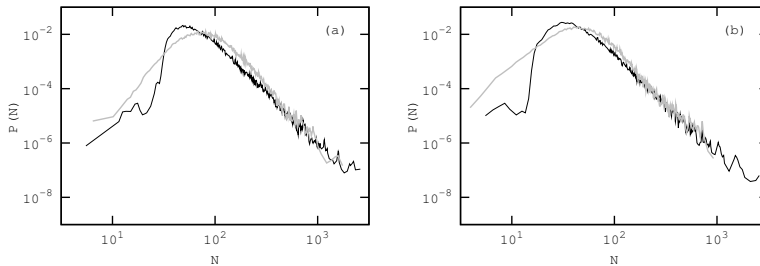


Fig. 6. Modeled, black curve, and empirical, gray curve, PDF of trading activity, N , for (a) IBM and (b) MMM stocks in the time interval $\tau_d = 300\text{s}$. Parameters are the same as in Fig. 3. $\tau_0 = 5\text{s}$ for IBM and $\tau_0 = 7.25\text{s}$ for MMM stocks.

the PSD are universal and arise from the scaled driving SDE (29) and (33). One can obtain the PSD of the selected stock's trading sequence by scaling model PSD, Figure 3 (b), by $1/\tau_0^2$. The PDF of integrated trading activity N is more sensitive to the market fluctuations. Even the intraday fluctuations of market activity, which are not included in this model, make considerable influence on PDF of N for low values. Nevertheless, as we demonstrate in Figure 6, the model is able to reproduce the power-law tails very well.

In this section we have shown results of the empirical analysis of stocks traded on NYSE. We have used those results as a basis for adjustment of the previously introduced trading activity model parameters. Aforementioned model is based on Poisson-like process, which we have introduced as scalable in previous sections, similar scalability as we see in this section is an inherent feature of actual financial markets.

A new form of scaled equations provides the universal description with the same parameters applicable for all stocks. The proposed new form of the continuous stochastic differential equation enabled us to reproduce the main statistical properties of the trading activity and waiting time, observable in the financial markets. In proposed model the fractured power-law distribution of spectral density with two different exponents arise. This is in agreement with the empirical power spectrum of the trading activity and volatility and implies that the market behavior may be dependent on the level of activity. One can observe at least two stages in market behavior: calm and excited. Ability to reproduce empirical PDF of inter-trade time

and trading activity as well as the power spectrum in very detail for various stocks provides a background for further stochastic modeling of volatility.

7. The stochastic model with a q -Gaussian PDF and power spectrum $S(f) \sim 1/f^\beta$

In section (3) we have introduced the class of SDE (12), (16) exhibiting power-law statistics and proposed Poisson like process modulated by this type of SDE. The latter serves as an appropriate model of trading activity in the financial markets Gontis et al. (2008). In this section we generalize the earlier proposed nonlinear SDE within the non-extensive statistical mechanics framework, Tsallis (2009), to reproduce the long-range statistics with a q -Gaussian PDF and power spectrum $S(f) \sim 1/f^\beta$.

The q -Gaussian PDF of stochastic variable r with variance σ_q^2 can be written as

$$P(r) = A_q \exp_q \left(-\frac{r^2}{(3-q)\sigma_q^2} \right), \tag{35}$$

here A_q is a constant of normalization, while q defines the power law part of the distribution. $P(r)$ is introduced through the variational principle applied to the generalized entropy Tsallis (2009), which is defined as

$$S_q = k \frac{1 - \int [p(r)]^q dr}{1 - q}.$$

The q -exponential of variable x is defined as

$$\exp_q(x) = (1 + (1 - q)x)^{\frac{1}{1-q}} \tag{36}$$

here we assume that the q -mean $\mu_q = 0$. With some transformation of parameters σ_q and q , namely

$$\lambda = \frac{2}{q-1}, \quad r_0 = \sigma_q \sqrt{\frac{3-q}{q-1}},$$

we can rewrite the q -Gaussian PDF in a more transparent form:

$$P_{r_0,\lambda}(r) = \frac{\Gamma(\lambda/2)}{\sqrt{\pi}r_0\Gamma(\lambda/2 - 1/2)} \left(\frac{r_0^2}{r_0^2 + r^2} \right)^{\frac{\lambda}{2}}. \tag{37}$$

Looking for the appropriate form of the SDE we start from the general case of a multiplicative equation in the Ito convention with Wiener process W :

$$dr = a(r)dt + b(r)dW. \tag{38}$$

If the stationary distribution of SDE (38) is the q -Gaussian (37), then the coefficients of drift, $a(r)$, and diffusion, $b(r)$, in the SDE are related as follows Gardiner (1986):

$$a(r) = -\frac{\lambda}{2} \frac{r}{r_0^2 + r^2} b(r)^2 + b(r) \frac{db(r)}{dr}. \tag{39}$$

From our previous experience modeling one-over- f noise and trading activity in financial markets Gontis & Kaulakys (2004); Kaulakys et al. (2005), building nonlinear stochastic differential equations exhibiting power law statistics Kaulakys et al. (2006); Kaulakys & Alaburda

(2009), described here in previous sections, we know that processes with power spectrum $S(f) \sim 1/f^\beta$ can be obtained using the multiplicative term $b(r) \sim r^\eta$ or even a slightly modified form $(r_0^2 + r^2)^{\frac{\eta}{2}}$. Therefore, we choose the term $b(r)$ as

$$b(r) = \sigma(r_0^2 + r^2)^{\frac{\eta}{2}} \quad (40)$$

and, consequently, by Eq. (39) we arrive at

$$a(r) = \sigma^2 \left(\eta - \frac{\lambda}{2} \right) (r_0^2 + r^2)^{\eta-1} r. \quad (41)$$

Having defined drift, Eq. (41), and diffusion, Eq. (40), terms one obtains this SDE:

$$dr = \sigma^2 \left(\eta - \frac{\lambda}{2} \right) (r_0^2 + r^2)^{\eta-1} r dt + \sigma(r_0^2 + r^2)^{\frac{\eta}{2}} dW. \quad (42)$$

Note that in the simple case $\eta = 1$ Eq. (42) coincides with the model presented in the article Queiros (2007) with

$$b(r) = \sqrt{\frac{\theta}{P(r)^{\frac{\lambda}{2}}}}, \quad a(r) = -\frac{\theta}{r_0^2} \left(\frac{\lambda}{2} - 1 \right) r \quad (43)$$

Further we will investigate higher values of η in order to cache long-range memory properties of the absolute return in the financial markets. First of all, let us scale our variables

$$x = \frac{r}{r_0}, \quad t_s = \sigma^2 r_0^{2(\eta-1)} t \quad (44)$$

to reduce the number of parameters and to get simplified equations. Then SDE

$$dx = \left(\eta - \frac{\lambda}{2} \right) (1 + x^2)^{\eta-1} x dt_s + (1 + x^2)^{\frac{\eta}{2}} dW_s \quad (45)$$

describes a stochastic process with a stationary q -Gaussian distribution

$$P_\lambda(x) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\lambda/2)}{\Gamma(\lambda/2 - 1/2)} \left(\frac{1}{1 + x^2} \right)^{\frac{\lambda}{2}} \quad (46)$$

and the power spectral density of the signal $S(f)$

$$S(f) = \frac{A}{f^\beta}, \quad \beta = 1 + \frac{\lambda - 3}{2(\eta - 1)} \quad (47)$$

$$A = \frac{(\lambda - 1)\Gamma(\beta - 1/2)}{2\sqrt{\pi}(\eta - 1)\sin(\pi\beta/2)} \left(\frac{2 + \lambda - 2\eta}{2\pi} \right)^{\beta-1} \quad (48)$$

with $0.5 < \beta < 2$, $4 - \eta < \lambda < 1 + 2\eta$ and $\eta > 1$. Eqs. (47-48) were first derived for the multiplicative point process in Gontis & Kaulakys (2004); Kaulakys et al. (2005) and generalized for the nonlinear SDE (42) in Kaulakys et al. (2006); Kaulakys & Alaburda (2009). Although Eq. (42) coincides with Eq. (15) in ref. Kaulakys & Alaburda (2009) only for high values of the variable $r \gg r_0$, these are the values responsible for the PSD. Note that the frequency f in equation (47) is the scaled frequency matching the scaled time t_s (44). The scaled equations (44)-(48) define a stochastic model with two parameters λ and η responsible for the power

law behavior of the signal PDF and power spectrum. Numerical calculations with Eq. (45) confirm analytical formulas (46-48) (see ref. Kaulakys & Alaburda (2009)).

We will need a more sophisticated version of the SDE to reproduce a stochastic process with a fractured PSD of the absolute return, which is observable in financial markets. Having in mind the statistics of the stochastic model (45) defined by Eqs. (46)-(48) and numerical modeling with more sophisticated versions of the SDE (27),(29), we propose an equation combining two powers of multiplicativity

$$dx = \left(\eta - \frac{\lambda}{2} - \left(\frac{x}{x_{max}} \right)^2 \right) \frac{(1+x^2)^{\eta-1}}{((1+x^2)^{\frac{1}{2}}\epsilon + 1)^2} x dt_s + \frac{(1+x^2)^{\frac{\eta}{2}}}{(1+x^2)^{\frac{1}{2}}\epsilon + 1} dW_s. \tag{49}$$

In modified SDE (49) model parameter ϵ divides area of x diffusion into two different power law regions to ensure the PSD of $|x|$ with two power law exponents. A similar procedure has been introduced in the model of trading activity Gontis et al. (2008), see previous sections. This procedure provides an approach to the market with behavior dependent on the level of activity and exhibiting two stages: calm and excited. Thus it is not surprising that Eq. (49) models the stochastic return x with two power law statistics, i.e., the PDF and the PSD, reproducing the empirical power law exponents of the trading return in the financial markets. At the same time, via the term $(\frac{x}{x_{max}})^2$ we introduce the exponential diffusion restriction for the high values of x as the markets in the excited stage operate on the limit of non-stationarity. One can solve Eq. (49) in the same manner we did solve trading activity related SDE (29) and (33). Thus we introduce the variable step of numerical integration

$$h_k = \kappa^2 \frac{((x_k^2 + 1)^{\frac{1}{2}}\epsilon + 1)^2}{(x_k^2 + 1)^{\eta-1}},$$

the differential equation (49) transforms into the set of difference equations

$$x_{k+1} = x_k + \kappa^2 \left(\eta - \frac{\lambda}{2} - (x_k^\epsilon)^\eta \right) x_k + \kappa (x_k^2 + 1)^{\frac{1}{2}} \epsilon_k \tag{50}$$

$$t_{k+1} = t_k + \kappa^2 \frac{((x_k^2 + 1)^{\frac{1}{2}}\epsilon + 1)^2}{(x_k^2 + 1)^{\eta-1}} \tag{51}$$

The continuous stochastic variable x does not include any time scale as the return defined in a time window τ should. Knowing that the return is an additive variable and depends on the number of transactions in a similar way to trading activity, we define the scaled return X in the time period τ as the integral of the continuous stochastic variable $X = \int_t^{t+\tau} x(t_s) / \tau dt_s$. Note that τ here is measured in scaled time units Eq. (44) though relation between model and empirical time series scales can be established and is very useful then adjusting model and empirical statistical properties.

It is worth recalling that integration of the signal in the time interval τ does not change the behavior of the power spectrum for the frequencies $f \ll \frac{1}{\tau}$. This is just the case we are interested in for the long-range memory analysis of financial variables and we can expect Eqs. (47-48) to work for the stochastic variable X as well.

We have also previously analyzed the influence of signal integration on the PDF in previous modeling of trading activity (see Gontis & Kaulakys (2004)). Integration of the nonlinear stochastic signal increases the exponent of the power law tails in the area of the highest values

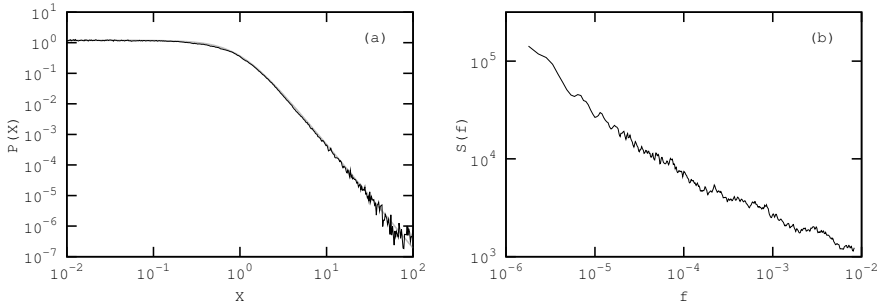


Fig. 7. (a) The numerically calculated PDF of $|X| = |\int_t^{t+\tau} x(t)/\tau dt_s|$ from Eq. (51) (black line), in comparison with the theoretical distribution $2P(x)$ Eq. (46) (gray line), and (b) the numerically calculated PSD of $|X|$. Model parameters are set as follows: $\eta = 5/2$, $\lambda = 3.6$, $\tau = 0.0001$ and $\epsilon = 0.01$.

of the integrated signal. This hides fractured behavior of the X PDF, which arises for x as a consequence of the two powers in the multiplicative term of Eq. (49).

In Fig. 12 we demonstrate (a) the numerically calculated PDF of $|X|$ in comparison with the theoretical distribution $2P(x)$ Eq. (46) and (b) the numerically calculated power spectrum of $|X|$ with parameters appropriate for reproducing statistics for the absolute return in financial markets.

8. The double stochastic model of return in financial market

Recently we proposed the double stochastic model of return in financial markets Gontis et al. (2010) based on the nonlinear SDE (49). The main advantage of proposed model is its ability to reproduce power spectral density of absolute return as well as long-term PDF of return. In the model proposed we assume that the empirical return r_t can be written as instantaneous q -Gaussian fluctuations ζ with a slowly diffusing parameter r_0 and constant $\lambda = 5$

$$r_t = \zeta \{r_0(t), \lambda\}. \tag{52}$$

q -Gaussian distribution of ζ defining the random instantaneous r_t can be written as follows:

$$P_{r_0, \lambda}(\zeta) = \frac{\Gamma(\frac{\lambda}{2})}{r_0 \sqrt{\pi} \Gamma(\frac{\lambda}{2} - \frac{1}{2})} \left(\frac{r_0^2}{r_0^2 + \zeta^2} \right)^{\lambda/2}, \tag{53}$$

with parameter $r_0(t)$ serving as a measure of instantaneous volatility of return fluctuations. We will model stochastic $r_0(t)$ in the similar way as trading activity in it's stochastic model. In this case nonlinear SDE (49) will serve as modulating one for q -Gaussian return fluctuations. The empirical evidence of this assumption is published in Gontis et al. (2010). The return, $|r|$, we define as absolute difference of logarithms of asset prices, p , in two different time moments separated by time interval τ :

$$r(t, \tau) = |\ln[p(t + \tau)] - \ln[p(t)]|. \tag{54}$$

In empirical analysis we consider dimensionless returns normalized by its dispersion calculated in the whole length of realization. It is worth to notice that $r(\tau)$ is an additive variable,

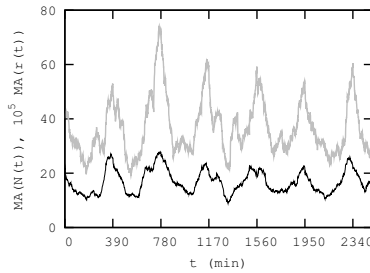


Fig. 8. An example of a moving average for 60 min of empirical absolute returns per minute (gray curve) in comparison with the corresponding moving average of trading activity, number of trades per minute (black curve). Scales are adjusted (magnitude of return time series multiplied by 10^5).

i.e., if $\tau = \sum_i \tau_i$, then $r(\tau) = \sum_i r(\tau_i)$, or in the continuous limit the sum may be replaced by integration. We do propose to model the measure of volatility r_0 by the scaled continuous stochastic variable x , having a meaning of average return per unit time interval. By the empirical analyses of high frequency trading data on NYSE Gontis et al. (2010) we introduced relation:

$$r_0(t, \tau) = 1 + \frac{\bar{r}_0}{\tau_s} \left| \int_{t_s}^{t_s + \tau_s} x(s) ds \right|, \tag{55}$$

where \bar{r}_0 is an empirical parameter and the average return per unit time interval $x(t_s)$ can be modeled by the nonlinear SDE (49), written in a scaled dimensionless time $t_s = \sigma_r^2 t$.

We have performed the empirical analyses (see Gontis et al. (2010)) of the tick by tick trades of 24 stocks, ABT, ADM, BMY, C, CVX, DOW, FNM, GE, GM, HD, IBM, JNJ, JPM, KO, LLY, MMM, MO, MOT, MRK, SLE, PFE, T, WMT, XOM, traded on the NYSE for 27 months from January, 2005, recorded in the Trades and Quotes database. We summed empirical tick by tick returns into one-minute returns to adjust the continuous stochastic model presented. Detailed analysis of the empirical data from the NYSE provides evidence that long-range memory properties of the return strongly depend on fluctuations of trading activity. In Fig. 8 we demonstrate strong correlation of the moving average of absolute returns per minute with the moving average of trading activities (number of trades per minute). Here for the empirical sequences of one-minute returns $\{r_t\}_{t=1}^T$ or trading activities $\{N_t\}_{t=1}^T$ we calculate moving averages MA defined as the centered means for a selected number of minutes n ; for example, $MA(r_t)$ is

$$MA(r_t) = \frac{1}{n} \sum_{j=t-n/2}^{t+n/2-1} r_j. \tag{56}$$

The best correlation can be achieved when the moving averages are calculated in the period from 60 to 100 minutes.

In order to account for the double stochastic nature of return fluctuations - a hidden slowly diffusing long-range memory process and rapid fluctuations of the instantaneous price changes - we decompose the empirical one-minute return series into two processes: the background

fluctuations and the high amplitude rapid fluctuations dependent on the first one modulating. To perform this empirical decomposition already presented as background idea of the model (52), we assume that the empirical return r_t can be written as instantaneous q -Gaussian fluctuations with a slowly diffusing parameter r_0 dependent on the moving average of the empirical return r_t :

$$r = \xi\{r_0(\text{MA}(r_t)), \lambda_2\}, \quad (57)$$

where $\xi\{r_0, \lambda_2\}$ is a q -Gaussian stochastic variable with the PDF defined by Eq. (53) (the parameter q is $q = 1 + 2/\lambda_2$). In Eq. (57) the parameter r_0 depends on the modulating moving average of returns, $\text{MA}(r_t)$, and the empirically defined power law exponent λ_2 . From the empirical time series of the one-minute returns r_t one can draw histograms of r corresponding to defined values of the moving average $\text{MA}(r_t)$. The q -Gaussian PDF is a good approximation to those histograms and the adjusted set of r_0 for selected values of $\text{MA}(r_t)$ gives an empirical definition of the function

$$r_0(\text{MA}(r_t)) = 1 + 2 \times |\text{MA}(r_t)|. \quad (58)$$

The q -Gaussians with $\lambda_2 = 5$ and linear function $r_0(|\text{MA}(r_t)|)$ (58) give a good approximation of r fluctuations for all stocks and values of modulating $\text{MA}(r_t)$. The long-term PDF of moving average $\text{MA}(r_t)$ can be approximated by a q -Gaussian with $\bar{r}_0 = 0.2$ and $\lambda = 3.6$. All these empirically defined parameters form the background for the stochastic model of the return in the financial market.

Consequently, we propose to model the long-range memory stochastic return $\text{MA}(r_t)$ by $X = \frac{\bar{r}_0}{\tau_s} \int_{t_s}^{t_s + \tau_s} x(s) / ds$, where x is a continuous stochastic variable defined by Eq. (49) and $\bar{r}_0 = \bar{r}_0 \times 2 = 0.4$. The remaining parameters ϵ , x_{max} and σ_t^2 can be adjusted for the best model fit to the empirical data and have values $\epsilon = 0.017$, $\sigma_t^2 = 1/3 \times 10^{-6} \text{s}^{-1}$ and $\tau_s = \tau \times \sigma_t^2 = 0.00002$; $x_{max} = 1000$.

The parameters of stochastic model were adjusted to the empirical tick by tick one minute returns. An excellent agreement between empirical and model PDF and power spectrum was achieved, see Fig 9 (a,b). Noticeable difference in theoretical and empirical PDFs for small values of return r are related with the prevailing prices of trades expressed in integer values of cents. We do not account for this discreteness in our continuous description. In the empirical PSD one-day resonance - the largest spike with higher harmonics - is present. This seasonality - an intraday activity pattern of the signal - is not included in the model either and this leads to the explicable difference from observed power spectrum.

9. Scaled comparison of model with empirical data

Seeking to discover the universal nature of financial markets we consider that all these parameters are universal for all stocks traded on various exchanges. To prove this we analyze empirical data from very different exchanges New York, one of the most developed with highly liquid stocks, and Vilnius, emerging one with stocks traded rarely. The comparison of model and empirical data scaling with increasing time window of return definition, τ , serves as very significant test for proposed stochastic description of financial markets.

Provided that we use scaled dimensionless equations derived while making very general assumptions, we expect that proposed model should work for various assets traded on different financial markets as well as for various time scales τ . We analyze tick by tick trades of 4 stocks, APG1L, PTR1L, SRS1L, UKB1L, traded on VSE for 50 months since May, 2005, trading data was collected and provided for us by VSE. Stocks traded on VSE in comparison with NYSE are less liquid - mean inter-trade time for analyzed stocks traded on VSE is 362 s, while for

stocks traded on NYSE mean inter-trade time equals 3.02 s. The difference in trading activity exceeds 100 times. This great difference is related with comparatively small number of traders and comparatively small companies participating in the emerging VSE market. Do these different markets have any statistical affinity is an essential question from the theoretical point of market modeling.

First of all we start with returns for very small time scales $\tau = 60$ s. For the VSE up to 95% of one minute trading time intervals elapse without any trade or price change. One can exclude these time intervals from the sequence calculating PDF of return. With such simple procedure calculated PDF of VSE empirical return overlaps with PDF of NYSE empirical return (see Fig 9 (a)).

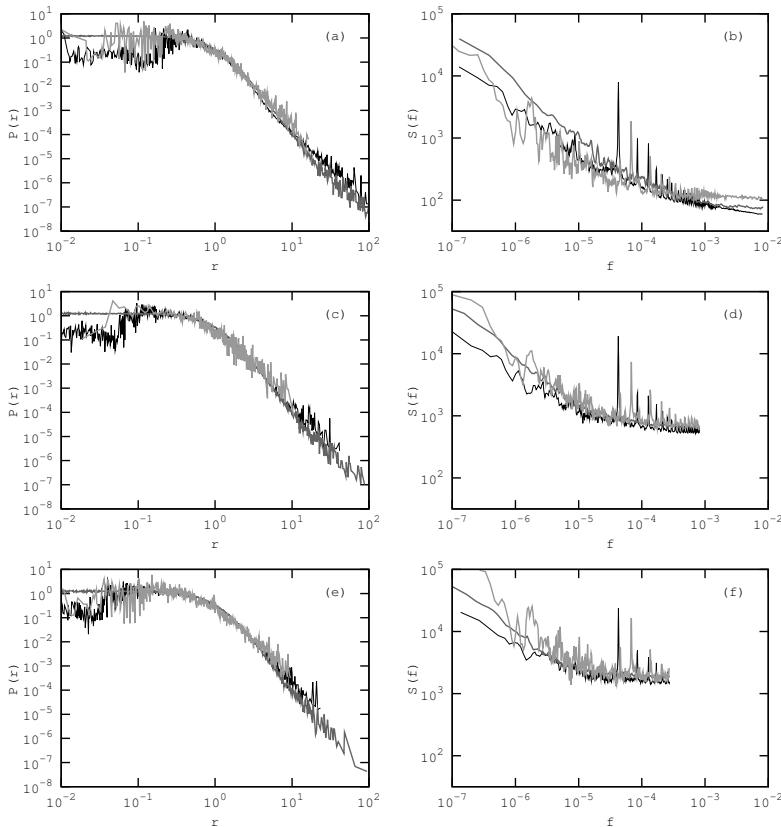


Fig. 9. Comparison of empirical statistics of absolute returns traded on the NYSE (black curves) and VSE (light gray curves) with model, defined by Eq. (49) and Eq. (55), statistics (gray curves). Model parameters are as follows: $\lambda = 5$; $\sigma_t^2 = 1/3 \cdot 10^{-6} s^{-1}$; $\lambda_0 = 3.6$; $\epsilon = 0.017$; $\eta = 2.5$; $\bar{r}_0 = 0.4$; $x_{max} = 1000$. PDF of normalized absolute returns is given on (a),(c),(e) and PSD on (b),(d),(f). (a) and (b) represents results with $\tau = 60$ s; (c) and (d) $\tau = 600$ s; (e) and (f) $\tau = 1800$ s. Empirical data from NYSE is averaged over 24 stocks and empirical data from VSE is averaged over 4 stocks.

One should use full time series of returns calculating the PSD. Nevertheless, despite low VSE liquidity, PSD of VSE and NYSE absolute returns almost overlap. Difference is clearly seen only for higher frequencies and smaller time windows, namely $\tau = 60$ s, and is related directly to the low VSE liquidity, which contributes to the white noise appearance. The different length of trading sessions in financial markets causes different positions of resonant intraday activity spikes. Thus one can conclude that even so marginal market as VSE retains essential statistical features as developed market on NYSE. At the first glance the statistical similarity should be even better for the higher values of return time scale τ .

Therefore further we investigate the behavior of returns on NYSE and VSE for increased values of $\tau = 60$ s and $\tau = 1800$ s with the specific interest to check whether proposed stochastic model scales in the same way as empirical data does. Apparently, as we can see in Fig 9 (d) and (f) PSDs of absolute returns on VSE and on NYSE overlap even better at larger time scale (600 seconds and 1800 seconds). This serves as an additional argument for the very general origin of long range memory properties observed in very different, liquidity-wise, markets. The nonlinear SDE is an applicable model to cache up observed empirical properties. PDFs of absolute return observed in both markets (see Fig 9 (c) and (e)) are practically identical, though we still have to ignore zero returns of VSE to arrive to the same normalization of PDF.

10. Conclusions

In the last sections we introduced a double stochastic process driven by the nonlinear scaled SDE ((49)) reproducing the main statistical properties of the absolute return, observed in the financial markets. Seven parameters of the model enable us to adjust it to the sophisticated power-law statistics of various stocks including long-range behavior. The scaled non-dimensional form of equations gives an opportunity to deal with averaged statistics of various stocks and compare behavior of different markets. All parameters introduced are recoverable from the empirical data and are responsible for the specific statistical features of real markets. Seeking to discover the universal nature of return statistics we have analysed and compared extremely different markets in New York and Vilnius and adjust the model parameters to match statistics of both markets. The most promising result of this research is discovered increasing coincidence of the model with empirical data from the New York and Vilnius markets and between markets, when the time scale of return τ is growing. Further analyses of empirical data and proposed model reasoning by agent behavior is ongoing.

11. References

- Beran, J. (1994). *Statistics for Long-Memory Processes*, Chapman and Hall, New York.
- Borland, L. & Bouchaud, J.-Ph. (2004). On a multi-timescale statistical feedback model for volatility fluctuations. E-print, arXiv:physics/0507073.
- Cont, R. (2005). Long range dependence in financial markets, In: *Fractals in Engineering*, Lut-ton, E. & Vehel, J., (Eds.), p. 159-180, Springer, Berlin.
- Gabaix, X.; Gopikrishnan, P.; Plerou, V. & Stanley, H.E. (2003). A theory of power-law distributions in financial market fluctuations. *Nature*, 423, p. 267-270.
- Gardiner, C.W. (1986). *Handbook of Stochastic Methods*, Springer, Berlin.
- Gontis, V. & Kaulakys, B. (2004). Multiplicative point process as a model of trading activity. *Physica A*, 343, p. 505-514. Gontis, V. & Kaulakys, B. (2004). Modeling financial markets by the multiplicative sequence of trades. *Physica A*, 344, p. 128-133. Gontis, V.;

- Kaulakys, B.; Alaburda, M. & Ruseckas J. (2004). Evolution of Complex Systems and $1/f$ Noise: from Physics to Financial Markets. *Solid State Phenomena*, 97-98, p. 65-70.
- Gontis, V. & Kaulakys, B. (2006). Long-range memory model of trading activity and volatility. *Journal of Statistical Mechanics*, P10016.
- Gontis, V. & Kaulakys, B. (2007). Modeling long-range memory trading activity by stochastic differential equations. *Physica A*, 382, p. 114-120.
- Gontis, V.; Kaulakys, B. & Ruseckas, J. (2008). Trading activity as driven Poisson process: comparison with empirical data. *Physica A*, 387, p. 3891-3896.
- Gontis, V.; Ruseckas, J. & Kononovičius A. (2010). A long-range memory stochastic model of the return in financial markets. *Physica A*, 389, p. 100-106.
- Kaulakys, B. & Meškauskas, T., (1998). Modelling $1/f$ noise. *Physical Review E*, 58, p. 7013-7019.
- Kaulakys, B. (1999). Autoregressive model of $1/f$ noise. *Physical Letters A*, 257, 1-2, p. 37-42.
- Kaulakys, B. (2000). On the intrinsic origin of $1/f$ noise. *Microelectronics Reliability*, 40, p. 1787-1790.
- Kaulakys, B.; Gontis, V. & Alaburda, M. (2005). Point process model of $1/f$ noise versus a sum of Lorentzians. *Physical Review E*, 71, 051105.
- Kaulakys, B.; Ruseckas, J.; Gontis, V. & Alaburda, M. (2006). Nonlinear stochastic models of $1/f$ noise and power-law distributions. *Physica A*, 365, p. 217-221.
- Kaulakys, B. & Alaburda, M. (2009). Modeling scaled processes and $1/f^\beta$ noise using non-linear stochastic differential equations. *Journal of Statistical Mechanics*, February 2009, P02051.
- Kirman, A. & Teyssiere, G. (2002). Microeconomic models for long-memory in the volatility of financial time series. *Studies in nonlinear dynamics and econometrics*, 5, p. 281-302.
- Li, W. (2009). URI: <http://www.nslj-genetics.org/wli/1fnoise>.
- Lux, T. & Marchesi, M. (2000). Volatility clustering in financial markets: a microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, 3, p. 675-702.
- Plerou, V.; Gopikrishnan, P.; Amaral L.; Gabaix, X. & Stanley, H.E. (2000). Economic fluctuations and anomalous diffusion. *Physical Review E*, 62, R3023-R3026.
- Plerou, V.; Gopikrishnan, P.; Gabaix, X. et al. (2001). Price fluctuations, market activity, and trading volume. *Quantitative Finance*, 1, p. 262-269.
- Duarte Queiros, S.M. (2007). On a generalised model for time-dependent variance with long-term memory. *Europhysics Letters*, 80, 30005
- Ruseckas, J. & Kaulakys B. (2010). $1/f$ noise from nonlinear stochastic differential equations. *Physical Review E*, 81, 031105.
- Sato, A.H. (2004). Explanation of power law behavior of autoregressive conditional duration processes based on the random multiplicative process. *Physical Review E*, 69, 047101.
- Takayasu, M. & Takayasu, H. (2003). Self-modulation processes and resulting generic $1/f$ fluctuations. *Physica A*, 324, p. 101-107.
- Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics*. Springer, ISBN 978-0-387-85358-1, New York.
- Willinger, W.; Taqqu, M. & Teverovsky, V. (1999). Stock market prices and long-range dependence. *Finance and Stochastics*, 3, p. 1-13.

