Analyzing memory effects of complex systems from time series

Daniel T. Schmitt and Michael Schulz
Theoretische Physik, Universität Ulm, D-89069 Ulm, Germany
(Received 12 June 2005; revised manuscript received 31 January 2006; published 18 May 2006)

A numerical algorithm is presented in order to determine all coefficients of the Mori-Zwanzig equation from a given finite time series. The algorithm is applicable to observables of arbitrary complex systems. Meteorological and financial systems are investigated. By analyzing directional variables and amplitudes we are able to observe and discuss memory effects on different time scales. We show that analyzing the memory kernel provides important insights into the dynamics of a complex system.

DOI: 10.1103/PhysRevE.73.056204 PACS number(s): 05.45.Tp, 87.23.Ge, 92.60.Gn

I. INTRODUCTION

Time series analysis [1,2] is an important tool used to characterize complex systems in physical research. Current investigations focus on rare events [3,4], the long-term behavior of correlation functions [5,6], and anomalous phenomena in the distribution of detrended fluctuations [7]. Through several numerical and analytical approaches in time series analysis, we strive to understand and classify climate evolution [8], river runoff and temperature records [9], medical electrocardiogram (ECG) and electroencephalogram (EEG) signals [10], financial markets [11,12], and DNA and protein sequences [13].

An important property of many *complex systems* is the large amount of degrees of freedom.[1] Examples range from granular material, and the earth climate, to man-made systems, such as traffic in urban environments, computer networks, and markets. Despite the fact that all objects involved in these systems obey known physical laws, for a variety of reasons, we are often unable to solve these problems. The reasons behind this are twofold. First, we cannot measure all degrees of freedom of most complex systems. Second, even if we were able to measure all degrees of freedom, in most cases, there are too many. Such that even the biggest super computer could not calculate their evolution. Interestingly enough, we are not even primarily interested in all degrees of freedom, but rather in only very few "relevant" observables.

Classical time series analysis often assumes a so-called component model for the relevant observables. In any component model, one assumes that there is a deterministic component and a stochastic component. The deterministic component often contains a monotone term named trend and periodic terms, which are often called seasonal or cyclic terms. The person who develops a specific component model empirically incorporates the available information about the system in the deterministic component. Therefore, the *principal of minimal information* [14] is applied to the stochastic component, which leads to the interpretation of noise. Consequently, the concurrent strength and weakness of classical time series analysis is that the deterministic component, the model, is empirically chosen based on the experience of the person committing the analysis. The subjective nature of this part of the analysis turns it into more of an art than science.

The *multiscale method* [15] is a well-established technique to systematically treat systems with slow and fast components. Our method however focuses on the separation of internal and external dynamics based on a projection formalism.

Mori-Zwanzig (MZ) theory [16–18] provides a projection formalism to obtain exact equations of motion for the relevant observables only. The dynamics of all other "irrelevant" degrees of freedom are hidden in so-called memory kernels and sometimes fast fluctuating residual forces. Thus, memory can be expected for any observable in a complex system where variables[2] are projected out. Although there are other projection formalisms [19–21] leading to different equations, the Mori-Zwanzig equation (MZE) is the only linear integrodifferential equation. Its functional structure is independent of the considered complex system.

In this paper, we present an algorithm to determine the empirical MZ equation from one finite time series of one relevant observable.[3] The fact that we do not need any information about the dynamics of the complex system, apart from approximate stationarity, makes it applicable to almost any complex system.

Together with the mastering technique, our method is capable of resolving the detailed memory kernel over a wide range of timescales. Compared to detrended fluctuation analysis (DFA) [22] we do not assume power-law correlations. On the other hand, DFA is more robust when nonstationarity is involved. The idea of a memory term is also common to the family of autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), fractal autoregressive integrated moving average (FARIMA), etc., models [23]. These models usually assume uncorrelated Gaussian noise. This makes them easier to implement, but inconsistent with the microscopic equations of motion, from which the MZ equation is derived.

II. MORI-ZWANZIG-THEORY

The Mori-Zwanzig theory [16] provides a projection formalism to obtain evolution equations of relevant observables only.

[1]There are also low-dimensional systems that exhibit complex behavior, i.e., Sinai billiard.

[2]In most cases the majority of the variables are projected out.

[3]Besides relaxation, the one-dimensional case also allows oscillationlike behavior.

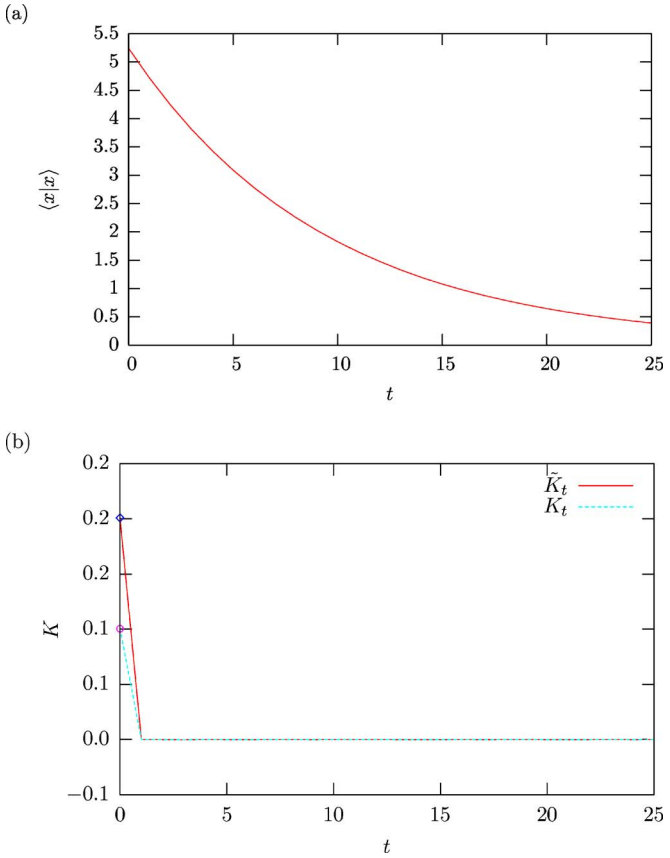

FIG. 1. (Color online) Ornstein-Uhlenbeck process. (a) The autocorrelation function, $\langle x|x\rangle$ decays exponentially. (b) The memory kernels K_t and $\tilde{K}_t$ show Markovian behavior. The frequency coefficient is $\Omega=\tilde{K}_0-K_0=0.10039$.

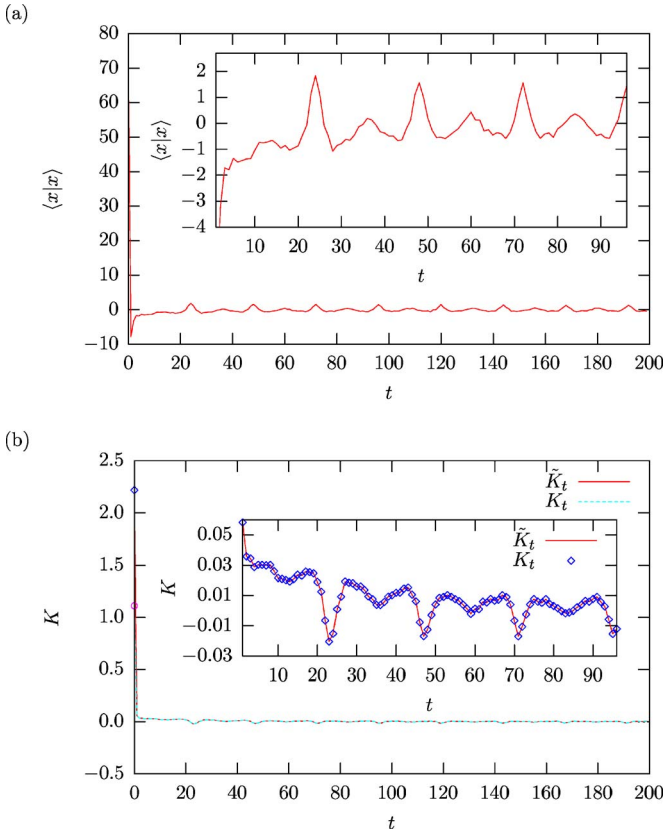

FIG. 2. (Color online) Changes in wind strength. (a) Autocorrelation function, $\langle x|x\rangle$. (b) Memory kernel, K_t and $\tilde{K}_t$, of the wind strength changes. The diamond at $t=0$ in the main plot is K_0 and the circle is $\tilde{K}_0$. $\Omega=\tilde{K}_0-K_0=1.11$. Both the autocorrelation function and the memory kernel observe clearly visible 12 h and, especially, 24 h periods. K_t and $\tilde{K}_t$ are calculated in two different ways (see Sec. III) and perfectly agree for $t\in[1,\ldots,n]$. This confirms relation (4) (see text for details).

Let $\{G\}$ be a set of relevant observables, $G_\alpha(t)$ $(\alpha=1,\ldots,M)$. All observables, $G_\alpha(t)$, are linearly independent, differentiable functions of the microscopic state $\Gamma=\{q_1,\ldots q_N,p1,\ldots,p_N\}$. Then, the time evolution of $G_\alpha(t)$ is given by the Mori-Zwanzig (MZ) equation

$$\frac{dG_\alpha(t)}{dt}=\sum_\gamma\left[\Omega_{\alpha\gamma}G_\gamma(t)-\int_{t_0}^{t}dt'\tilde{K}_{\alpha\gamma}(t-t')G_\gamma(t')\right]+f_\alpha(t,t_0). \quad (1)$$

The most general form of the memory kernel is $\tilde{K}_{\alpha\gamma}(t,t',t_0)$, where the system was last observed at the initial time t_0. The dependence of t and t' reduces to $t-t'$ if the system is embedded in a time-independent environment. Without loss of generality, Eq. (1) becomes autonomous[4] since the microscopic equations of motion that it was derived from can always be made autonomous by choosing a large enough system.

In Eq. (1), we further assume that the complex system is stationary. This is reasonable following the principle of Occam's Razor since it is impossible to decide if it is stationary or not based on one finite realization of the system. Because of stationarity the t_0 dependence in the memory kernel vanishes.

Although the MZ equation (1) is still exact, only in very special and simple systems one can derive the coefficients from the microscopic equations of motion (for example, Ref. [24]). As an approximation one substitutes the rest force terms, $f_\alpha(t,t_0)$, with a stochastic process that obeys Eqs. (2) and (4). That is why $f_\alpha(t,t_0)$ are often called *noise terms*.

At the point in time, t_0, there is full and secure information about $G_\alpha(t_0)$; thus, we know the initial condition of the relevant observables $G_\alpha(t_0)=G_{0\alpha}$. It is worth mentioning that, in general, the rest force terms, $f_\alpha(t,t_0)$, are completely different for different t_0 since the microscopic initial conditions vary as well (even if $G_{0\alpha}$ happens to be equal).

By adding a constant to $G_\alpha(t)$, one can always make the ensemble average over the noise term vanish

$$\langle f_\alpha(t,t_0)\rangle_{t_0}=0. \quad (2)$$

The following identities apply:

[4]Autonomous means that an equation does not depend on time-dependent external forces or fields or time-dependent boundary conditions.

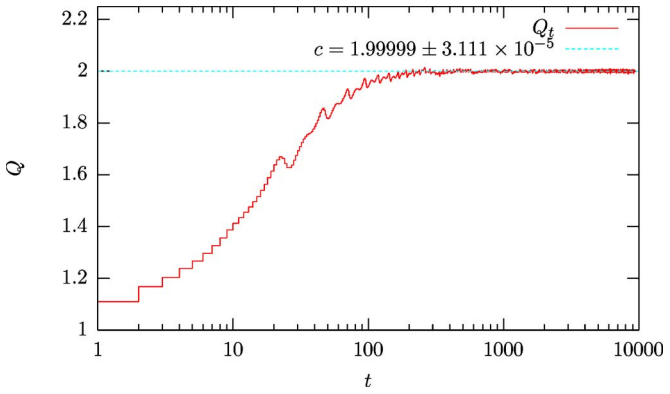

FIG. 3. (Color online) Cumulated memory kernel Q_t of the wind strength changes. The horizontal line was fitted to Q_t in the interval $t \in [500, 1000]$ The elementary time unit is 1 h.

$$\langle G_\beta(t_0) | f_\alpha(t,t_0) \rangle_{t_0} = 0 \tag{3}$$

$$\sum_\beta H^{-1}_{\gamma\beta} \langle f_\beta(t_0,t_0) | f_\alpha(t,t_0) \rangle_{t_0} = \tilde{K}_{\alpha\gamma}(t), \tag{4}$$

where $H_{\gamma\beta} = \langle G_\gamma(t_0) | G_\beta(t_0) \rangle_{t_0}$ (see Ref. [25]).

On the other hand, identity (4) can be derived solely from the functional structure of Eq. (1) and the assumption that Eq. (3) holds. In this case, it is not necessary to use any projection formalism. We remark that linear models with memory but without identity (4) were presented in the literature [26].

Because of the identity (3) it is possible to obtain a similar equation for the autocorrelation function, $c_{G_\alpha G_\beta}(t) = \langle G_\alpha(t) | G_\beta(t_0) \rangle$

$$\frac{dc_{G_\alpha G_\beta}(t)}{dt} = \sum_\gamma \left[\Omega_{\alpha\gamma} c_{G_\gamma G_\beta}(t) - \int_{t_0}^{t} dt' \tilde{K}_{\alpha\gamma}(t-t') c_{G_\gamma G_\beta}(t') \right]. \tag{5}$$

In case of only one relevant observable, Eq. (5) simplifies to

$$\frac{dc_{GG}(t)}{dt} = \Omega c_{GG}(t) - \int_{t_0}^{t} dt' \tilde{K}(t-t') c_{GG}(t'). \tag{6}$$

The problem now is to determine the coefficients, Ω, $\tilde{K}(t-t')$, and $f(t,t_0)$ from one finite time series.

III. ALGORITHM

In this section, the aim is to develop a numerical algorithm that determines the coefficients of the MZ equation, given a time series x_t of a relevant observable under the assumption of *stationarity*. The time series is assumed to be discrete in time. This is no limitation, since measured time series are always discrete and sometimes may approximate a continuous process.

Equation (6) serves as a convenient starting point since the noise term, $f_\alpha(t)$, was eliminated through averaging. Numerically, the autocorrelation function $c_{GG}(t)$ can be estimated in the following way:

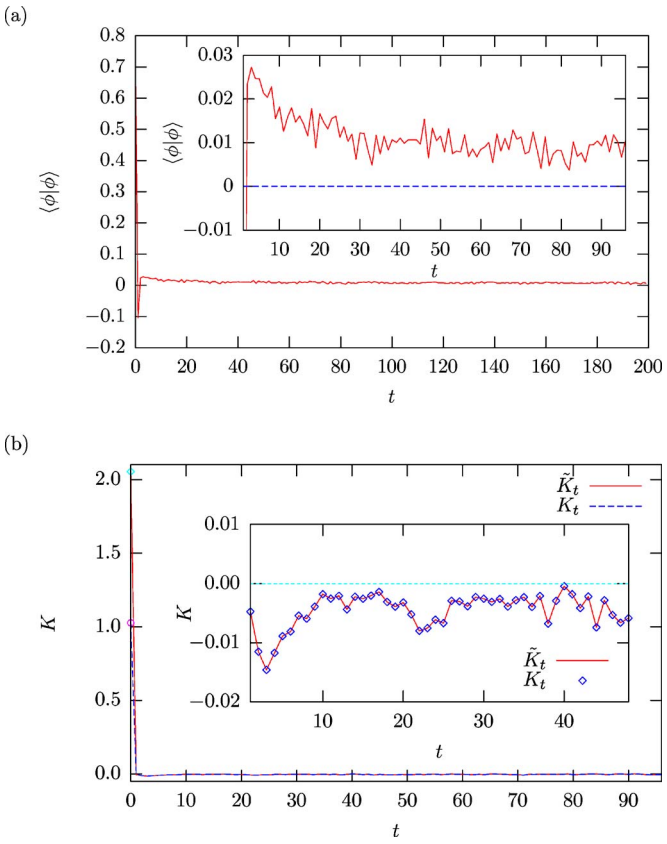

FIG. 4. (Color online) Changes in wind direction, ϕ. (a) Autocorrelation function, $\langle \phi | \phi \rangle$. (b) Memory kernels K_t and $\tilde{K}_t$, of the wind direction changes. The diamond at $t=0$ in the main plot is K_0 and the circle is $\tilde{K}_0$. $\Omega = \tilde{K}_0 - K_0 = 1.131$. The elementary time unit is 1 h.

$$c_{GG}[t] = \frac{1}{N-t} \sum_{i=0}^{N-t-1} (x_i - \langle x \rangle)(x_{i+t} - \langle x \rangle),$$

where we used the assumption of stationarity.[5]

In order to numerically treat Eq. (6), we choose a corresponding time-ordered discrete version

$$c_{GG}[t] - c_{GG}[t-1] = -\sum_{i=0}^{t-1} K_{(t-i-1)} c_{GG}[i], \tag{7}$$

where Ω is part of K_0. Equation (7) can easily be solved for K_t.

In order to separate the frequency Ω from K_0 and to check whether the memory kernel K_t was determined consistently, we use the relation (4).

The MZ equation in a time-ordered discrete form is given by

[5]Stationarity in the sense that ensemble average is equal to the time average.

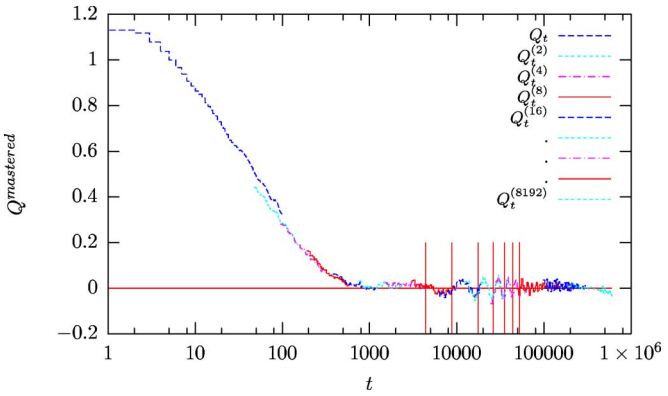

FIG. 5. (Color online) Cumulated mastered memory kernel Q_t^{mastered} of the wind direction changes. The vertical lines indicate $1,2,\ldots,4$ yr. The master technique is explained in Sec. IV B. The elementary time unit is 1 h.

$$x[t] - x[t-1] = -\sum_{i=t_0}^{t-1} K_{(t-i-1)}x[i] + f_{t_0}[t-1]. \quad (8)$$

Therefore, $f_{t_0}[t-t_0]$ can be calculated from the known time series, $x[t]$ in the following way:

$$f_{t_0}[t-t_0] = x[t-t_0+1] - x[t-t_0] + \sum_{i=t_0}^{t-t_0} K_{(t-t_0-i)}x[i]. \quad (9)$$

It is important to point out that t_0 is the point in time where the value of the relevant observable is known. Even though stationarity of the time series x_t ensures that one can substitute ensemble averaging by time averaging we still have a completely different noise term, $f_{t_0}[t-t_0]$, for every initial time t_0.

The discrete version of relation (4) then reads

$$\tilde{K}_t = \frac{1}{\langle x^2\rangle(0)}\langle f_{t_0}[t_0]|f_{t_0}[t-t_0]\rangle_{t_0}, \quad (10)$$

where $\langle x^2\rangle(0)$ is the variance of the time series and $\langle\cdots\rangle_{t_0}$ denotes that the average is taken over t_0.

Ideally, K_t and $\tilde{K}_t$ are equal except for the first element K_0 and $\tilde{K}_0$. The frequency Ω can be obtained from

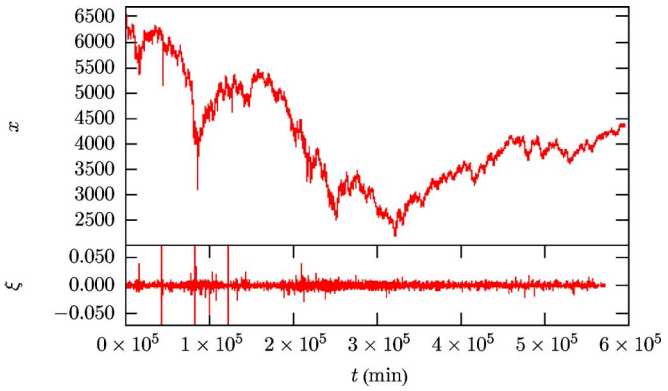

FIG. 6. (Color online) DAX time series. Minute data from 13:59 16.02.2001 to 19:37 21.02.2005.

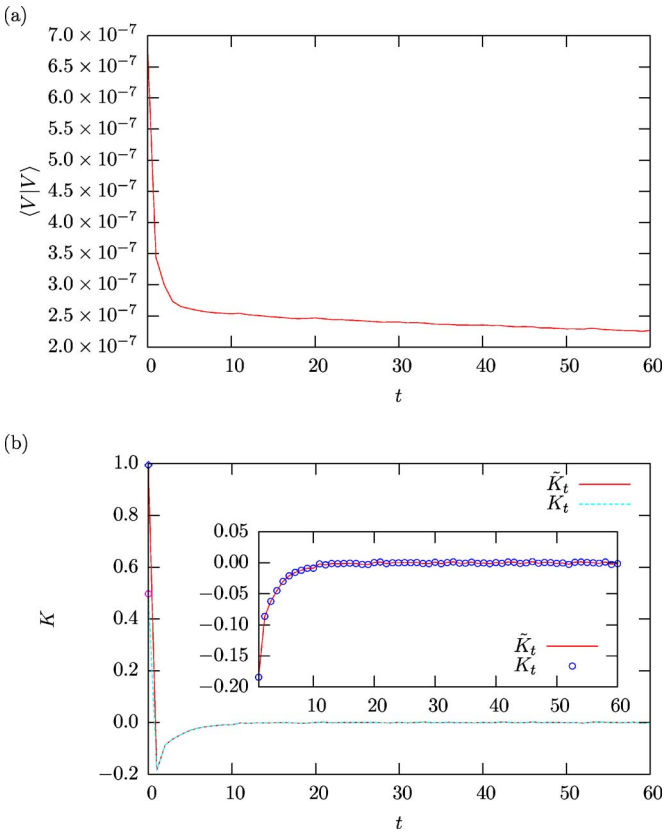

FIG. 7. (Color online) DAX volatility. (a) Autocorrelation function, $\langle V|V\rangle$. (b) Memory kernels K_t and $\tilde{K}_t$. The diamond at $t=0$ in the main plot is K_0 and the circle is $\tilde{K}_0$. $\Omega=\tilde{K}_0-K_0=0.497$. The elementary time unit is 1 min.

$$\Omega = \tilde{K}_0 - K_0. \quad (11)$$

IV. APPLICATIONS/RESULTS

A. Artificial data

In this instance, we use an Ornstein-Uhlenbeck (OU) process (12) to check our algorithm

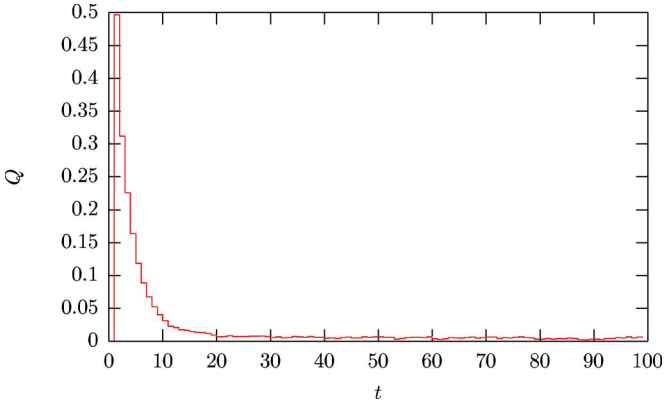

FIG. 8. (Color online) Cumulated memory kernel Q_t of the DAX volatility. The elementary time unit is 1 min.

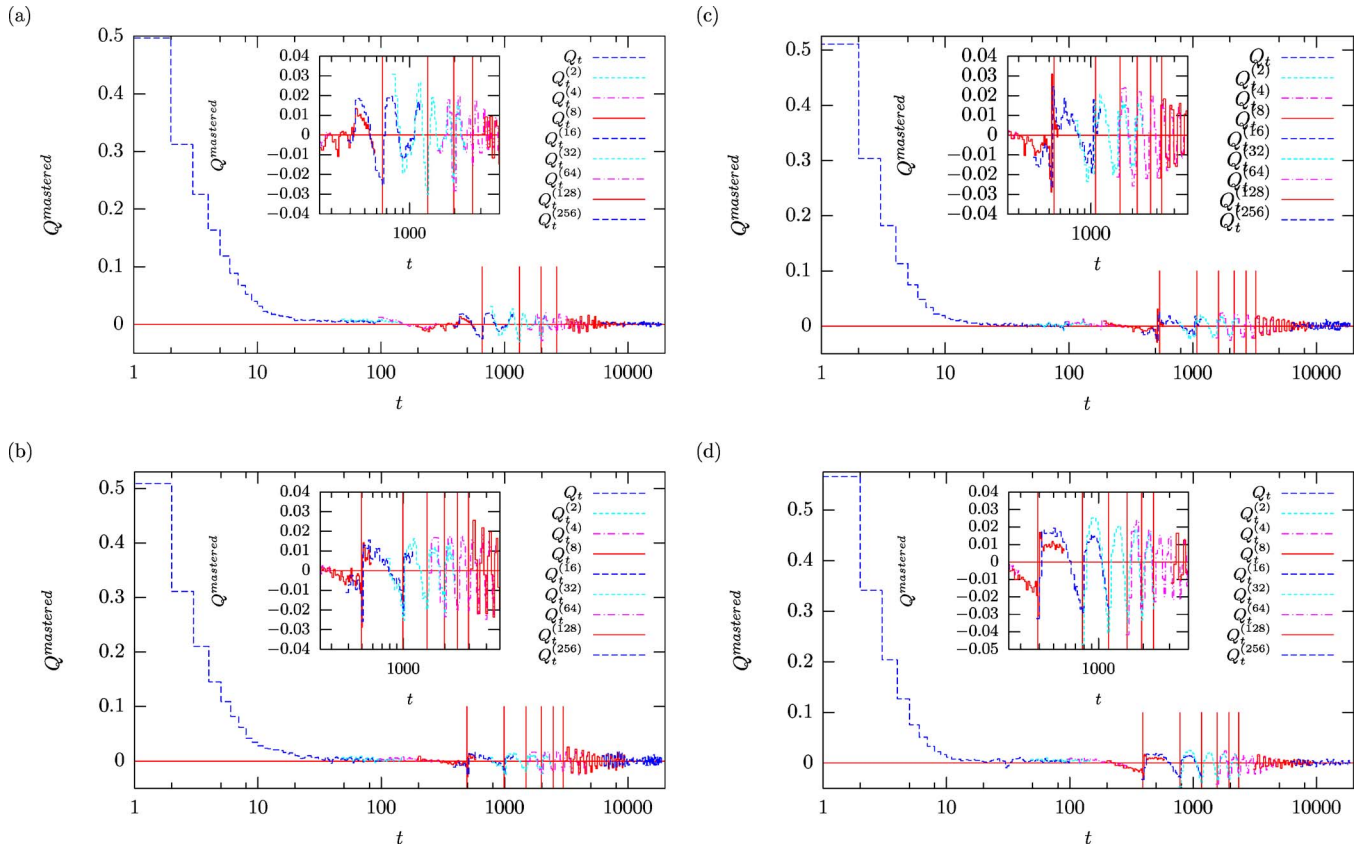

FIG. 9. (Color online) Master plots of cumulated memory kernels of the absolute logarithmic price changes V form different stock indices. Insets show a close-up of the oscillations. Vertical lines indicate 1, 2, 3, and 4 days. (a) Mastered cumulated memory kernel Q_t of the DAX. (b) Mastered cumulated memory kernel Q_t of the MICEX. (c) Mastered cumulated memory kernel Q_t of the FTSE. (d) Mastered cumulated memory kernel Q_t of the NASDAQ. Note that hours traded per day are different for each stock exchange. The master technique is analog to the one used in Fig. 5 (see text for details). The elementary time unit is 1 min.

$$\frac{dx(t)}{dt} = \Omega_{OU}x(t) + W(t). \tag{12}$$

We generate 500 000 data points with $\Omega_{OU}=-0.1$ and a standard Gaussian distributed $W(t)$. In terms of the MZ equation, the equation reads

$$\frac{dx(t)}{dt} = \Omega x(t) - \int_0^t \tilde{K}(t-t')x(t')dt' + W(t)$$

$$= -\int_0^t K(t-t')x(t')dt' + W(t). \tag{13}$$

Since the asymptotic variance is $\langle x^2\rangle = -\frac{1}{2\Omega_{OU}}$, we expect $K_0 = 0.1$ and $\tilde{K}_0 = \frac{1}{\langle x^2\rangle}\langle W^2\rangle = 2\times\Omega_{OU}\times 1 = 0.2$. This is in perfect agreement with Fig. 1(b).

Figure 1(a) shows the exponential decay of the autocorrelation function. In Fig. 1(b) the memory kernels K_t and $\tilde{K}_t$ are plotted. The frequency matrix in Eq. (13) is correctly estimated, $\Omega=0.100\,39$, which leads to $\Omega_{OU}=-0.099\,61$.

B. Physical time series

In this section, we analyze wind data since the atmosphere is a macroscopic system with a high degree of complexity. Our observables are wind strength v and wind direction ψ at Potsdam, Germany, starting from January 1, 1893 to April 30, 1999 for every hour of the day [27]. This data set was analyzed in Ref. [28] to find logarithmic corrections to the mean square displacement, considering the wind vector to describe a two-dimensional random walk.

We abstain from any form of detrending since our aim is to analyze the memory kernel of the original data. Let us analyze the memory kernel K_t of the changes in wind strength, x_t

$$x_t = v_t - v_{t-1}. \tag{14}$$

In Fig. 2(a), the autocorrelation function of x_t is shown. Figure 2(b) shows the corresponding memory kernels, K_t and $\tilde{K}_t$. It is important to point out that even though K_t and $\tilde{K}_t$ have been calculated in two different ways (see Sec. III) they perfectly agree for $t\in[1,\ldots,200]$. Thus, the relation (4) holds.

Here, we used a stochastic integrodifferential equation with memory of type (1) to describe the dynamics of a relevant observable, namely, wind strength changes. Assuming identity (3), we verified the relation (4) as a necessary consequence.

In Fig. 2(b) 12 h and, in particular, 24 h periods are clearly visible. Since the memory kernel K_t contributes nega-

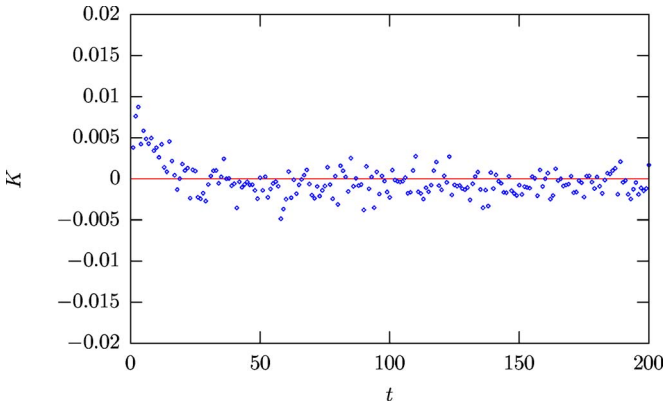

FIG. 10. (Color online) Short-term memory kernel K_t of the sgn(ξ) of the DAX. The elementary time unit is 1 min.

tively to changes of x_t, certain changes in wind strength make it more likely to experience the same changes, 12 h and, particularly, 24 h later. Similar periods are obtained by analyzing the wind strength instead of the change in wind strength.

In Fig. 3, the cumulated memory kernel $Q_t=\Sigma_{i=0}^{t}K_i$ of the changes in wind strength are plotted. After ~240 h or 10 days the cumulated memory kernel saturates, which means that after ~10 days, the wind strength changes show Markovian behavior. The saturation level of ~2 h^{-1} can be interpreted as an effective viscous friction in the system as in an Ornstein-Uhlenbeck process. As a result of friction, wind speeds, on average, do not get stronger over time as one would expect from a diffusive process.

Next we investigate wind direction changes. If we assume that the wind does not change more than π within 1 h, we can define the wind direction changes according to

$$\alpha = \psi_t - \psi_{t-1}$$

$$\phi_t = \begin{cases} \alpha - 2\pi, & \alpha > \pi \\ \alpha + 2\pi, & \alpha < -\pi \\ \alpha, & \text{otherwise.} \end{cases} \tag{15}$$

In Fig. 4, the autocorrelation function, $\langle \phi | \phi \rangle$, and the memory kernel K_t of the wind direction changes is plotted.

In contrast to the memory kernel of changes in wind strength, up to 96 h, there is little periodic structure in the memory kernel for wind direction changes. There is a slight overshooting into the negative domain, which then slowly creeps back to zero.

If we calculate the memory kernel K_t of the changes in wind direction for larger time lags t the result is very noisy. Smoothing and filter techniques do not produce convincing results. Thus, in order to investigate long-term memory we use the master technique, which is a standard procedure used to match observables on different time scales [29].

We coarse grain the changes in wind direction ϕ_t by taking the arithmetic average of adjacent nonoverlapping windows of length n. These new values, $\overline{\phi}_{n\Delta t}$, are now the average wind direction change in this time window. The fact that the averaging windows do not overlap is important in order to not introduce artificial correlations. Next we calculate the memory kernel, $K_t^{(n)}$, from $\overline{\phi}_{n\Delta t}$ for $n \in \{1,2,4,8,16,\ldots,2048\}$. Then plot the cumulated memory kernel, $Q_t^{(n)}=\Sigma_{i=0}^{t}K_i^{(n)}$ into one master plot (Fig. 5). For more clearly arranged plots, the beginning and end part of $Q_t^{(n)}$ $n \in \{2,4,\ldots,2048\}$ are omitted.

The mastered cumulated memory, $Q_t^{(n)}=\Sigma_{i=0}^{t}K_i^{(n)}$, of the wind direction changes in Fig. 5 decays to a first minimum at ~1536 h. At ~1 yr, the negative value of Q^{mastered} indicates a repulsive memory effect. The oscillations can be explained through the seasonal dependence of the dominating wind direction.

C. Financial time series

In this section we analyze minute data from the DAX (Fig. 6).[6] Let x_t be the closing prices of each 1 min time interval. Periods in which no trading took place are cut out. As a result, overnight intervals or holidays, for example, are also treated as a 1 min bin. We found that the following results are robust to the exclusion of overnight intervals or the exclusion of afterhour trading data. Next, we define the common logarithmic price changes defined in

$$\xi_t = \ln\frac{x_t}{x_{t-1}} = \ln x_t - \ln x_{t-1}. \tag{16}$$

The correlation time of logarithmic price changes ξ_t in liquid markets has decreased over the last centuries [30] and is usually in the order of the time resolution. Thus, in terms of the *efficient market hypothesis*, existing correlations would immediately be exploited by trades to earn money and would therefore vanish. It is usually impossible to resolve a detailed structure in the autocorrelation function, $\langle \xi | \xi \rangle$, of logarithmic price changes. Therefore, we analyze the absolute value, $V=|\xi_t|$, as a measure of volatility and the sign, sgn(ξ_t), of the logarithmic price changes ξ_t, separately,

$$\xi_t = |\xi_t|\text{sgn}(\xi_t). \tag{17}$$

We defined the sgn(x) function in a completely symmetric way

$$\text{sgn}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases} \tag{18}$$

Using the algorithm described above, we can find the memory kernel and the frequency of the MZ equation, and in this way, describe the evolution of the absolute logarithmic price changes V. Figure 7 shows the autocorrelation function and the memory kernel of the absolute logarithmic price changes on the minute time scale. It is well known that the autocorrelation function of the volatility is long-range correlated [31]. We see long-range correlations as well, but here we focus on the analysis of the memory kernel in the moderate time regime.

In Fig. 7(b), we use the difference between $\tilde{K}_0$ and K_0 to determine the frequency, $\Omega=0.496\,782\,1$. Besides for $t=0$,

[6]Data obtained from http://www.fin-rus.com/analysis/export/-eng-/default.asp [32].

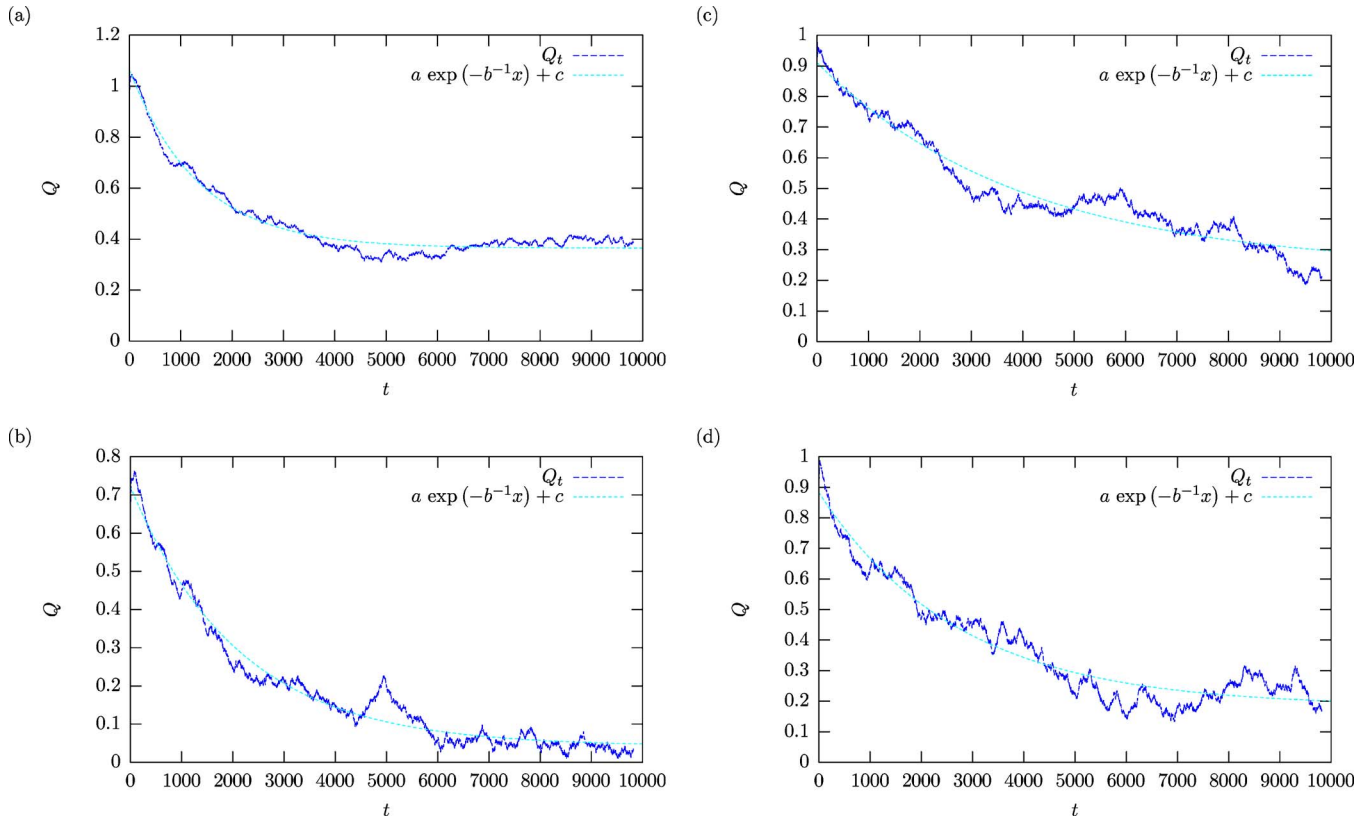

FIG. 11. (Color online) Cumulated memory Q_t of the sgn(ξ) from different stock indices. In order to estimate a typical time scale where memory effects exist, an exponential function was fitted. The elementary time unit is 1 min. (a) Cumulated memory of the sgn(ξ) of the DAX, $b=1350$ min=2.04 days. (b) Cumulated memory of the sgn(ξ) of the MICEX, $b=2114$ min=4.27 days. (c) Cumulated memory of the sgn(ξ) of the FTSE, $b=3997$ min=7.4 days. (d) Cumulated memory of the sgn(ξ) of the NASDAQ, $b=2712$ min=6.95 days. Note that hours traded per day are different for each stock exchange.

K_t and $\tilde{K}_t$ cannot be distinguished. The absolute square difference of K_t and $\tilde{K}_t$, $t \in 1,2,\ldots,100$, is 1.85×10^{-10}. This shows again that relation (4) holds our algorithm works consistently.

When the memory kernel K_t is very small for times larger than 20 min, erratic fluctuations make it hard to decide if there is a significant difference from zero. Hence, analyzing the cumulated memory kernel Q_t is helpful since systematic differences from zero accumulate. Furthermore, the cumulated memory kernel can be interpreted as a viscous friction coefficient in the sense of a *Markov approximation* of the MZ equation.

Figure 8 shows the cumulated memory kernel. Interestingly, the cumulative memory kernel almost vanishes after ~30 min, meaning that if we look at longer time scales, V seems to diffuse freely.

In order to obtain the memory kernel for larger time lags we again use the master technique discussed in Sec. IV B. The masterplots (Fig. 9) contain the cumulated memory kernels, $Q_t^{(n)}$ for $n \in \{1,2,4,8,16,32,64,128,256\}$.

The master technique works very well for the volatility of a variety of different stock indices (Fig. 9). Generally, one observes that the cumulated memory kernel decays to almost zero within half an hour. Oscillations then start again on the daily timescale. Negative peaks indicating driving effects occur for time lags of multiples of one trading day. The pattern between one and two trading days is very similar to the pattern between two and three trading days, etc. However, daily patterns do differ from stock index to stock index. For very large time lags, the envelope of these oscillations becomes narrower. Hence, the system "forgets."

Let us now investigate the sign of the logarithmic price changes, sgn(ξ). Looking at the memory kernel of the sgn(ξ) from the DAX in Fig. 10, one clearly recognizes that the first 17 data points are above zero. The price dynamic seems to "remember" the sign of the price change for at least 17 min. For larger time lags, the impression that most of the points are below zero (repulsive memory) can more clearly be seen when K_t is accumulated [Fig. 11(a)].

To investigate longer-lasting memory effects in the dynamics of the sign, we look at the accumulated memory kernel, $Q_t = \Sigma_{i=0}^{t} K_i$ in Fig. 11. All accumulated memory kernels decay nicely. In order to estimate a rough time scale, we fitted them with an exponential decay. Surprisingly, we find that apart from the short time behavior (~17 min) discussed above memory effects of 2–7 days are present in the dynamics of the sign.

V. CONCLUSIONS

In this paper, we presented an algorithm to empirically determine all components of a Mori-Zwanzig equation for

one observable of a complex system from one time series. For a complex system, this is usually the only accessible realization of the process. The generalization of the algorithm to the multivariate case is always possible and will be presented in a subsequent paper.

The algorithm was applied to wind data and the financial data of major stock indices. In both cases we analyzed directional variables and amplitudes. The directional variable in the case of the wind data is direction change and in the case of the financial data, the sign of the price fluctuation. The amplitude in the case of the wind data is the change in wind strength. In the case of the financial data, it is the volatility.

The memory kernel of the wind strength changes shows very distinct half-day and, especially, daily periods. The accumulated memory suggests a finite generalized dissipative force, which prevents diffusive behavior and, therefore, larger and larger wind speeds over time.

The memory kernel of the wind direction changes has little periodic structure on the daily time scale. The mastered accumulated memory reveals seasonal structure. We find the daily wind dynamics in the wind strength and seasonal dynamics in the wind direction. Qualitatively, this demonstrates the everyday experience of stronger winds in the mornings and evenings, more westerly winds in the summer, and easterly winds in the winter in central Europe.

For financial time series, the memory kernel of the volatility is pronounced at short time scales ($\sim$10 min). For intermediate time lags (up to 0.5 trading days), it is approximately zero. In contrast, we observe daily periods up to 10 trading days characterizing the long-term behavior of the memory. The fact that the daily oscillations decay for longer time lags indicates that the numerical truncating error and the finite length of the time series have not yet taken over.

Another surprising result is that apart from the fast short-time decay of the memory kernel of the signs, we find an additional long-range decay in the order of several trading days. As expected, shuffling of the original data sets results in vanishing memory effects. The results for the financial data were found to be robust to the exclusion of overnight intervals and the exclusion of afterhour trading data.

In principle, the presented algorithm can be applied to a wide range of time series. We demonstrated that the analysis of the memory kernels is an alternative method apart from the standard analysis of correlation functions, which often leads to different insight into the dynamics of complex systems.

ACKNOWLEDGMENTS

We thank the Deutscher Wetterdienst for the data. We thank FINAM Investment Holding for the data.

[1] H. Abarbanel, *Analysis of Observed Chaotic Data* (Springer-Verlag, Berlin, 1996).
[2] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis* (Cambridge University Press, Cambridge, England, 1997).
[3] A. Bunde, J. F. Eichner, J. W. Kantelhardt, and S. Havlin, Phys. Rev. Lett. **94**, 048701 (2005).
[4] D. R. Cox, J. R. Stat. Soc. Ser. B. Methodol. **34**, 187 (1972).
[5] M. Raberto, E. Scalas, G. Cuniberti, and M. Riani, Physica A **269**, 148 (1999).
[6] P. Cizeau, Y. Liu, M. Meyer, C.-K. Peng, and H. E. Stanley, Physica A **245**, 441 (1997).
[7] J. W. Kantelhardt, S. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley, Physica A **316**, 87 (2002).
[8] A. Bunde, S. Havlin, E. Koscielny-Bunde, and H.-J. Schellnhuber, Physica A **302**, 255 (2001).
[9] A. Bunde, J. Eichner, S. Havlin, and J. Kantelhardt, Physica A **342**, 308 (2004).
[10] Y. Yamamoto and R. L. Hughson, J. Appl. Physiol. **71**, 1143 (1991).
[11] V. Plerou, P. Gopikrishnan, and H. E. Stanley, Phys. Rev. E **71**, 046131 (2005).
[12] B. Podobnik, P. C. Ivanov, V. Jazbinsek, Z. Trontelj, H. E. Stanley, and I. Grosse, Phys. Rev. E **71**, 025104(R) (2005).
[13] P. Bernaola-Galvan, I. Grosse, P. Carpena, J. L. Oliver, R. Roman-Roldan, and H. E. Stanley, Phys. Rev. Lett. **85**, 1342 (2000).
[14] H. Haken, *Information and Self-Organization: A Macroscopic Approach to Complex Systems*, 2nd ed. (Springer, New York, 2000).
[15] W. E and B. Engquist, Not. Am. Math. Soc. **50**, 1062 (2003).
[16] R. Zwanzig, Phys. Rev. **124**, 983 (1961).
[17] H. Mori, Prog. Theor. Phys. **33**, 423 (1965).
[18] H. Mori, Prog. Theor. Phys. **34**, 399 (1965).
[19] B. Robertson, Phys. Rev. **144**, 151 (1966).
[20] B. Robertson, Phys. Rev. **160**, 175 (1967).
[21] B. Robertson, Phys. Rev. **166**, 206 (1968).
[22] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, Physica A **295**, 441 (2001).
[23] C. Chatfield, *Time Series Forecasting* (CRC Press, Boca Raton, FL, 2000).
[24] M. Einax and M. Schulz, J. Chem. Phys. **115**, 2282 (2001).
[25] E. Fick and G. Sauermann, *Quantenstatistik Dynamischer Prozesse* (Verlag Hary Deutsch-Thun, Frankfurt, 1986).
[26] M. Potters, J. P. Bouchaud, and L. Laloux, e-print physics/0507111.
[27] For further information see the Deutscher Wetterdienst website, http://www.dwd.de
[28] B. M. Schulz, M. Schulz, and S. Trimper, Phys. Lett. A **291**, 87 (2001).
[29] E. Donth, *Relaxation and Thermodynamics in Polymers* (Akademie-Verlag, Berlin, 1992).
[30] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing* (Cambridge University Press, Cambridge, England, 2003), Chap. 6.2.2.
[31] Y. Liu, P. Gopikrishnan, P. Cizeau, M. Meyer, C.-K. Peng, and H. E. Stanley, Phys. Rev. E **60**, 1390 (1999).
[32] For further information see the FINAM Investment Company website, http://www.fin-rus.com/analysis/export/-eng-/default.asp