# Dynamical modelling of superstatistical complex systems

Erik Van der Straeten *, Christian Beck

*Queen Mary University of London, School of Mathematical Sciences, Mile End Road, London E1 4NS, UK*

## ARTICLE INFO

## ABSTRACT

We show how to construct the optimum superstatistical dynamical model for a given experimentally measured time series. For this purpose we generalise the superstatistics concept and study a Langevin equation with a memory kernel whose parameters fluctuate on a large time scale. We show how to construct a synthetic dynamical model with the same invariant density and correlation function as the experimental data. As a main example we apply our method to velocity time series measured in high-Reynolds-number turbulent Taylor–Couette flow, but the method can be applied to many other complex systems in a similar way.

## 1. Introduction

Many complex systems in nature exhibit a rich structure of dynamics, described by a mixture of different stochastic processes on various time scales. It has become common to call such dynamical processes with time scale separation *superstatistical* [1–8]. A fascinating feature is that many highly nonlinear equations, for example the Navier–Stokes equations, do effectively produce stochastic processes that can be very well approximated by linear stochastic differential equations with fluctuating parameters on a large time scale, i.e., superstatistical processes. Examples where superstatistical techniques have proved to be a powerful tool are Lagrangian trajectories of tracer particles in fully developed turbulence [2,3], share price dynamics with volatility fluctuations [4,7,8], atmospheric and geophysical processes [9], random matrix theory [10] and random networks [11]. The basic idea underlying most of these models is that locally the system is described by a simple Gaussian stochastic process, and only on a large scale does the mixture of various Gaussian processes with fluctuating variance lead to more complex behaviour.

So far the superstatistical model building has mainly concentrated on reproducing the correct stationary density of the nonlinear system under consideration. However, it is certainly important to incorporate dynamical information from the set of all higher-order correlation functions. In fact, experimentalists are often able to measure two relevant quantities: the stationary probability distribution of a fluctuating time series produced by the complex system under consideration, and the two-point correlation function for relatively small time scale differences. The purpose of this article is to describe how to construct the optimum superstatistical model for a given complex system taking these two experimentally measured quantities as input. As a main example, we will construct a suitable superstatistical model for measured velocity time series in turbulent Taylor–Couette flow, which is in excellent agreement with the experimental data. Our techniques presented here are quite generally applicable to large classes of complex systems where the local dynamics is well approximated by a generalised Langevin equation with memory kernel.

## 2. The standard Langevin equation

We start from a simple example that will be subsequently generalised. Consider locally a linear stochastic differential equation (without memory kernel) of the form

---

* Corresponding author.
  *E-mail addresses:* e.straeten@qmul.ac.uk (E. Van der Straeten), c.beck@qmul.ac.uk (C. Beck).

$$\frac{d}{dt}v(t) = -\gamma v(t) + \sqrt{\frac{2\gamma}{\beta}}g(t),\tag{1}$$

where $\gamma$ and $\beta$ are constants and $g(t)$ is Gaussian white noise. This generates the Ornstein–Uhlenbeck process, a Gaussian Markov process, with stationary distribution $p(v)$ and correlation function $c_v(\Delta)$:

$$p(v) = \sqrt{\frac{\beta}{2\pi}}e^{-\frac{1}{2}\beta v^2}$$

$$c_v(\Delta) = \lim_{t\to\infty}\langle v(t)v(t+\Delta)\rangle = \frac{1}{\beta}e^{-\gamma\Delta}.\tag{2}$$

$\langle . \rangle$ denotes an ensemble average. In the superstatistical approach, one regards the parameter $\beta$ as a random variable that fluctuates on a large time scale. If this random variable has the probability distribution $f(\beta)$, then the stationary (marginal) distribution of the complex system under consideration is given by

$$P(v) = \int_0^\infty d\beta f(\beta)\sqrt{\frac{\beta}{2\pi}}e^{-\frac{1}{2}\beta v^2}.\tag{3}$$

All kinds of distributions $P(v)$ can be generated in this way, for example power law decays in $v$ and stretched exponentials, depending on the properties of $f(\beta)$ [5]. In fact, an experimentally measured $P(v)$ can be used to construct $f(\beta)$ from Eq. (3). The superstatistical correlation function is given by

$$C_v(\Delta) = \int_0^\infty d\beta f(\beta)c_v(\Delta) = e^{-\gamma\Delta}\int_0^\infty d\beta f(\beta)\frac{1}{\beta} \propto c_v(\Delta).\tag{4}$$

Clearly, an experimentalist may ask whether such a simple model is appropriate for a given experimental data set. In most cases it will not be, because the measured decay of the correlation function is not simply given by $\exp(-\gamma\Delta)$ as is the case for the Ornstein–Uhlenbeck process. In fact, experimental data often show local correlation functions that are well described by a sum or difference of two exponentials, or sometimes there is oscillating behaviour. Our aim in the following is thus to extend the local dynamics in such a way that this more refined local dynamics, reproducing the correct (measured) correlation function, can be incorporated into the superstatistical modelling approach.

## 3. The generalised Langevin equation

A more general starting point is to consider locally a set of coupled linear stochastic differential equations

$$\frac{d}{dt}v_j(t) = -\sum_k \Theta_{j,k}v_k(t) + \Xi_j g_j(t).\tag{5}$$

The elements of the matrix $\Theta$ and the vector $\Xi$ are constants. The noise terms are denoted by $g_j(t)$. The idea of the theory of Mori and Zwanzig [12,13] is that some of the variables $v_j(t)$ are irrelevant, or inaccessible to the experimentalist. By eliminating these variables from the equations of motion, one ends up with a set of *non-Markovian* equations in the relevant variables only. In this article we study in detail a two-dimensional version of this problem, namely

$$\frac{d}{dt}v(t) = -\gamma v(t) + V(t) + \alpha g(t)\tag{6}$$

$$\frac{d}{dt}V(t) = -\eta V(t) + \sigma v(t).\tag{7}$$

Here, $\eta$ and $\gamma$ are positive constants. The constants $\alpha$ and $\sigma$ can have either sign. $g(t)$ denotes Gaussian white noise, i.e., $\langle g(t)\rangle = 0$ and $\langle g(t)g(t')\rangle = \delta(t-t')$. For the above dynamics, the two eigenvalues of the matrix $\Theta$ are $A \pm B$, with

$$2A = \eta + \gamma \quad \text{and} \quad 2B = \sqrt{(\eta-\gamma)^2 + 4\sigma}.\tag{8}$$

It is easy to see that the real parts of these eigenvalues are always positive provided $\eta\gamma - \sigma > 0$. In the remaining part of this article, we assume that this inequality holds. This ensures that the system will locally reach a stationary state [12].

Our physical interpretation is as follows. The velocity variable $v(t)$ decays with relaxation constant $\gamma$ and is subjected to two forces: a rapidly fluctuating force $g(t)$ that drives the dynamics and another force $V(t)$ that decays with relaxation constant $\eta$. This force $V(t)$, however, is not independent of the velocity $v(t)$. Rather, the change of $V(t)$ is proportional to the velocity $v(t)$, with proportionality constant $\sigma$. The above linear coupling scheme is one of the simplest, but still non-trivial, examples of coupled local dynamics between pairs of variables. Now assume that an experimentalist can only measure one

of these two variables. We chose $v(t)$ to be this variable and proceed by eliminating $V(t)$ from the set of equations (6), (7). In this way one obtains an equation in $v(t)$ with a memory kernel [12,13]:

$$\frac{\mathrm{d}}{\mathrm{d}t}v(t) = -\gamma v(t) + \sigma \int_0^t \mathrm{d}t' \mathrm{e}^{-\eta(t-t')}v(t') + \alpha g(t). \tag{9}$$

To obtain this formula, we ignored the contribution $V(0)\exp(-\eta t)$ because this term is unimportant for long times. Expression (9) contains an exponential memory kernel and is known in the literature as the generalised Langevin equation [12,13]. Usually, one assumes that the time dependence of the memory kernel and the correlation function of the stochastic term are equal. This equality is called the fluctuation-dissipation relation of the second kind [14] and the stochastic term $g(t)$ is then interpreted as internal noise. However, for the example studied in the present paper, the aforementioned time dependencies are not equal. Then, the stochastic term $g(t)$ is interpreted as external noise. Such noise terms have also been studied recently in the context of stochastic resonance [15] and anomalous diffusion [16].

Two well-known special cases are included in our local model. Taking the limit $\eta \to \infty$ in (9) yields the standard Langevin equation (1) without memory kernel. Equivalently, this case also arises on putting $\sigma = 0$, i.e. no coupling between velocity and force. Moreover, expressions (6), (7) reduce to the equations of motion of a Brownian harmonic oscillator for $\eta = 0$ and $\sigma < 0$. The variables $v(t)$ and $-V(t)$ are then interpreted as the velocity and the position of the oscillator, respectively.

## 4. The local model: the stationary distribution

Before implementing superstatistics, we first have to solve the above local model. Let us start with calculating the local stationary distribution $p(v)$ of the relevant variable $v(t)$. The Fokker–Planck equation [17] for the joint probability distribution $p(v, V, t)$ is

$$0 = \left[ \frac{\partial}{\partial t} + \frac{\partial}{\partial v}(V - \gamma v) + \frac{\partial}{\partial V}(\sigma v - \eta V) - \frac{\alpha^2}{2}\frac{\partial^2}{\partial v^2} \right] p(v, V, t). \tag{10}$$

The stationary distribution $p(v, V)$ is obtained by setting $\partial p(v, V, t)/\partial t = 0$ and is given by

$$p(v, V) \sim \exp\left[ -\frac{\gamma + \eta}{\sigma^2 \alpha^2}[(\sigma v - \eta V)^2 + (\eta \gamma - \sigma)V^2] \right]. \tag{11}$$

For the marginal distribution $p(v)$ we integrate out the dependence on $V$:

$$p(v) = \int_{-\infty}^{+\infty} \mathrm{d}V p(v, V) = \sqrt{\frac{\beta}{2\pi}}\mathrm{e}^{-\frac{1}{2}\beta v^2}, \tag{12}$$

where the parameter $\beta$ is given by

$$\beta = \frac{2(\gamma + \eta)(\eta \gamma - \sigma)}{\alpha^2(\eta \gamma - \sigma + \eta^2)}. \tag{13}$$

We see, as required for superstatistics, that the local stationary distribution is a Gaussian distribution.

## 5. The local model: the correlation function

To proceed further into a dynamical setting, we have to calculate a formal solution for $v(t)$. Let us denote the Laplace transform $\hat{f}(w)$ of a function $f(t)$ by

$$\hat{f}(w) = \int_0^\infty \mathrm{d}t \mathrm{e}^{-wt}f(t). \tag{14}$$
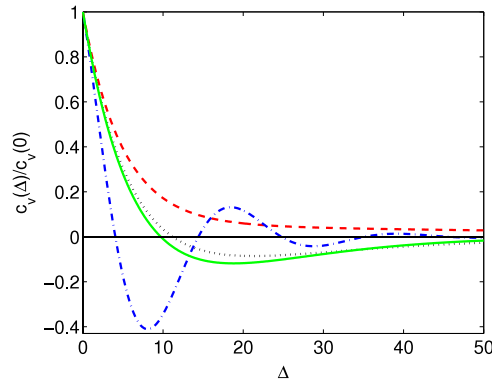
From Eqs. (6), (7) one obtains $\hat{v}(w)$ as

$$\hat{v}(w) = \hat{\varrho}(w)\left[ \frac{V(0)}{w + \eta} + v(0) + \alpha \hat{g}(w) \right], \tag{15}$$

where the function $\hat{\varrho}(w)$ is

$$\hat{\varrho}(w) = (w + \eta)(w + A + B)^{-1}(w + A - B)^{-1} \tag{16}$$

(see (8) for the definition of $A$ and $B$). By calculating the inverse Laplace transform of (15), one obtains

$$v(t) = v(0)\varrho(t) + V(0)\int_0^t \mathrm{d}t' \varrho(t - t')\mathrm{e}^{-\eta t'} + \alpha \int_0^t \mathrm{d}t' \varrho(t - t')g(t'). \tag{17}$$

**Fig. 1.** (Colour online) Plot of the correlation functions for different values of $\sigma$. The solid line is a representation of expression (21), i.e. the case $B = 0$. The dashed–dotted curve shows an oscillatory behaviour (22). The dashed and dotted curves are respectively a sum and a difference of two exponential functions (23).

The stationary correlation function $c_v(\Delta)$ is obtained by evaluating the ensemble average of the product of $v(t)$ and $v(t + \Delta)$ for $t \to \infty$. All the terms involving the initial conditions $v(0)$ and $V(0)$ that contribute to $c_v(\Delta)$ are vanishing for $t \to \infty$ and in this limit one ends up with the following expression for the correlation function:

$$c_v(\Delta) = \alpha^2 \int_0^\infty d\tau \varrho(\tau)\varrho(\tau + \Delta). \tag{18}$$

In order to obtain an explicit expression for $c_v(\Delta)$, we have to calculate the inverse Laplace transform of $\hat{\varrho}(w)$ given by (16). The result of this transformation depends on whether $B = 0$ or $B \neq 0$. However it is easy to prove that the following inequalities hold: $\mathscr{R}(A \pm B) > 0$ given $\eta > 0$, $\gamma > 0$ and $\eta\gamma - \sigma > 0$. This means that the Laplace transform is well-defined provided that the latter inequalities are satisfied. Under the assumption that $B = 0$, one obtains the following expression for the function $\varrho(t)$:

$$\varrho(t) = [1 + (\eta - A)t]e^{-At}, \tag{19}$$

while for the case $B \neq 0$ the result is

$$\varrho(t) = \frac{1}{2B}(-\eta + A + B)e^{-(A+B)t} - \frac{1}{2B}(-\eta + A - B)e^{-(A-B)t}. \tag{20}$$

We proceed by evaluating (18) to calculate an explicit expression for the correlation function. Under the assumption that $B = 0$, one obtains

$$\frac{c_v(\Delta)}{c_v(0)} = \frac{e^{-A\Delta}}{\eta^2 + A^2}[\eta^2 + A^2 + A(\eta^2 - A^2)\Delta]. \tag{21}$$

We split the case $B \neq 0$ into two subcases, depending on whether $B$ is imaginary or real. In the case where $B \neq 0$ and $\mathscr{R}(B) = 0$, the result for $c_v(\Delta)$ is

$$\frac{c_v(\Delta)}{c_v(0)} = e^{-A\Delta}[\Gamma' \sin(B'\Delta) + \cos(B'\Delta)], \tag{22}$$

while in the case where $B \neq 0$ and $\mathscr{R}(B) = B$, one ends up with

$$\frac{c_v(\Delta)}{c_v(0)} = \frac{e^{-A\Delta}}{1 + \Gamma}[e^{-B\Delta} + \Gamma e^{B\Delta}], \tag{23}$$

with $B' = B/i$ and

$$\Gamma' = \frac{A}{B'}\frac{\eta^2 - A^2 - B'^2}{\eta^2 + A^2 + B'^2}, \qquad \Gamma = \frac{B + A}{B - A}\frac{(B - A)^2 - \eta^2}{(B + A)^2 - \eta^2}. \tag{24}$$

It is easy to show that the signs of $\Gamma$ and $\sigma$ are equal. This means that the parameter $\sigma$ determines whether one ends up with a sum or a difference of exponentials in expression (23). As a consequence, there are four qualitatively different kinds of correlation functions contained in the set of stochastic differential equations (6), (7). This is illustrated in Fig. 1. The figure shows the correlation function for four different values of the coupling parameter $\sigma$ but with constant values of the other parameters $\gamma = 0.2$ and $\eta = 0.02$. Our chosen example values of $\sigma$ are $-0.0081$ (solid line), $-0.1$ (dashed–dotted line), $0.001$ (dashed line) and $-0.005$ (dotted line). Experimentally measured correlation functions in turbulence are often of strikingly similar shape; see for example Figure 2a in Ref. [18].
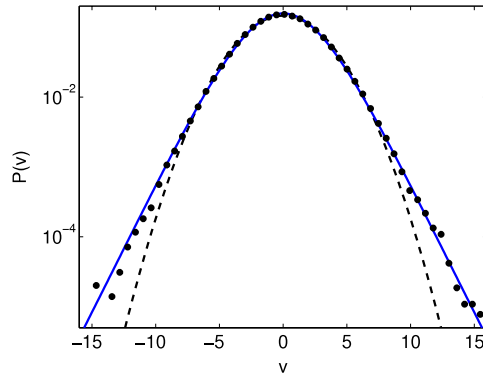
**Fig. 2.** (Colour online) Plot of the measured distribution $P(v)$ (● ● ●). The dashed line represents a Gaussian approximation $0.15\exp(-0.07v^2)$ for $P(v)$. The solid line is the superstatistical distribution (3), with $f(\beta)$ a log-normal distribution (25).

## 6. Superstatistics

Finally let us apply superstatistics to this local model by introducing large-scale fluctuations in $\beta$. We assume that $\beta$, as given by Eq. (13), is distributed according to some probability density $f(\beta)$ and that the changes of $\beta$ occur on a time scale that is much larger than the relaxation time of the correlation function $c_v(\Delta)$. Generally, this results in a superstatistical distribution $P(v)$ given by (3) that is non-Gaussian. The time dependence of the superstatistical correlation function $C_v(\Delta)$ given by (4) is equal to the time dependence of $c_v(\Delta)$ as long as the $\beta$-fluctuations are produced by fluctuations of $\alpha$ only.

We are now in a position to formulate the central result of this article. Let a complex system be given that exhibits time scale separation and superstatistical behaviour (see Ref. [3] for a test). Suppose we know $P(v)$ and $C_v(\Delta)$ from experimental measurements. How should we construct the optimum synthetic superstatistical model that is consistent with the measured data? Of course many nonlinear models are possible, but if one assumes that there is a local linear coupling between pairs of dynamical variables and that locally the stochastic process considered exhibits Gaussian behaviour, then the model (6), (7) with superstatistical parameter fluctuations is the most appropriate one. The distribution $f(\beta)$ of the parameter $\beta$ can be determined by comparing Eq. (3) with the experimentally measured distribution $P(v)$. Then, depending on the shape of the measured correlation function, the values of the parameters $\gamma$, $\eta$ and $\sigma$ can be determined.

To illustrate our general method, we study as an example time series obtained in an experiment performed by Lewis and Swinney [19,2]. The data set contains the values of a single velocity component $v(t)$ as a function of time $t$ in turbulent Taylor–Couette flow for Reynolds number $Re = 540.000$. The stationary probability distribution $P(v)$ exhibits non-Gaussian behaviour; see Fig. 2. A suitable choice for $f(\beta)$, motivated by the K62 theory of turbulence [20], is a log-normal distribution

$$f(\beta) = \frac{1}{b_1\sqrt{2\pi}\,\beta}\exp(-(\ln\beta - b_2)^2/2b_1^2). \tag{25}$$

Choosing $b_1 = 0.42$ and $b_2 = -1.91$ one obtains an excellent fit of the experimentally measured histogram; see Fig. 2. But there are of course many different stochastic processes generating the same stationary distribution $P(v)$. We can now construct a more precise synthetic model of the dynamics by taking into account the information contained in the measured correlation function. The experimentally measured correlation function $C_v(\Delta)/C_v(0)$ is shown in Fig. 3(a), together with a single exponential function $\exp(-0.030\Delta)$ and a sum of two exponentials; see expression (23) with $A = 0.024$; $B = 0.019$; $\Gamma = 0.571$. Fig. 3(a) clearly shows that a single exponential function is not able to properly fit the measured correlation function, while a sum of two exponentials yields an excellent fit. This means that our superstatistical memory-kernel approach is a suitable way to model these data.

In this paper, we did not consider the $\beta$-process. That is to say, we did not construct an explicit model for the time development of the slowly varying variable $\beta$. However, it is clear that for very long time scales the details of this process will influence the correlation function $C_v(\Delta)/C_v(0)$ and the stationary distribution $P(v)$. To investigate the behaviour for larger time scales $\Delta$, in Fig. 3(b) we show the measured and modelled correlation function in a double-logarithmic plot, which emphasises the long-time behaviour. We see that our model (solid line) fits the measured data well up to time scales $\Delta \approx 400$. Beyond that, there are large statistical fluctuations and the behaviour is dominated by details of the $\beta$-process. For example, an asymptotic power law decay could be constructed by considering a $\beta$-process whose long-term correlations decay with a power law. Taking the $\beta$-process into account goes beyond the scope of the current paper but is an interesting topic for further research. Our current approach as presented in this paper gives testable predictions for small and intermediate time scales $\Delta$. The asymptotic $\Delta \to \infty$ behaviour is then dominated by the chosen $\beta$-process.

Let us emphasise that obtaining a sum (and not a difference) of two exponentials for the correlation function (23) is non-trivial. For this it is crucial that $\sigma > 0$ and as a consequence $\eta \neq 0$ and $\gamma \neq 0$. Such correlation functions are also observed in biological time series for trajectories of motile cells [21] and have been recently obtained theoretically for two-dimensional stochastic motion with uncorrelated fluctuations of the speed and the direction of the motion [22].
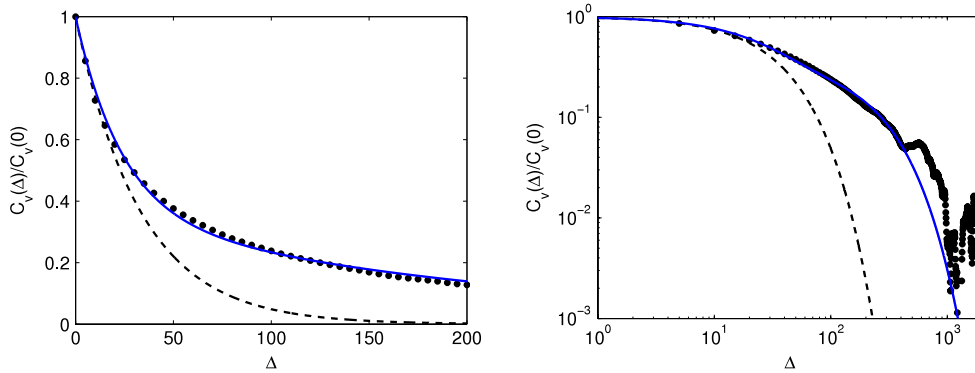
**Fig. 3.** (Colour online) Plot of the measured normalised correlation function $C_v(\Delta)/C_v(0)$ (● ● ●). The dashed and solid lines represent an approximation for $C_v(\Delta)/C_v(0)$ in terms of a single exponential $\exp(-0.030\Delta)$ and a sum of two exponentials (23), respectively. (a) Linear plot; (b) double-logarithmic plot.

## 7. Discussion

To summarise, in this article we have shown how to construct an optimum superstatistical dynamical model that exhibits the same invariant density and correlation function as an experimentally measured time series extracted from a complex system. The relevant universality class of complex systems to which our approach is applicable consists of systems where locally pairs of dynamical variables are coupled in the simplest possible (linear) way, as described by Eqs. (6), (7). This leads to a Langevin equation with memory kernel whose parameters fluctuate on a large time scale in a superstatistical way. As an example, we applied our approach to experimentally measured velocity time series in turbulent Taylor–Couette flow, obtaining excellent agreement with the experimental data. Our approach is quite generally applicable to many complex systems with time scale separation, as long as high-precision experimental data on the stationary density and correlation function are available.

## References

- [1] C. Beck, E.G.D. Cohen, Physica A 322 (2003) 267.
- [2] C. Beck, E.G.D. Cohen, H.L. Swinney, Phys. Rev. E 72 (2005) 056133.
- [3] E. Van der Straeten, C. Beck, Phys. Rev. E 80 (2009) 036108.
- [4] A. Reynolds, Phys. Rev. Lett. 91 (2003) 084503.
- [5] H. Touchette, C. Beck, Phys. Rev. E 71 (2005) 016131.
- [6] C. Beck, Phys. Rev. Lett. 98 (2007) 064502.
- [7] P. Jizba, H. Kleinert, Phys. Rev. E 78 (2008) 031122.
- [8] C. Anteneodo, S.M. Duarte Queiros, J. Stat. Mech. (2009) P10023.
- [9] A. Porporato, G. Vico, P.A. Fay, Geophys. Res. Lett. 33 (2006) L15402.
- [10] A.Y. Abul-Magd, et al., Phys. Rev. E 77 (2008) 046202.
- [11] S. Abe, S. Thurner, Phys. Rev. E 72 (2005) 036102.
- [12] R. Zwanzig, Nonequilibrium Statistical Mechanics, Oxford University Press, New York, 2000.
- [13] H. Mori, Progr. Theoret. Phys. 33 (1965) 423.
- [14] R. Kubo, Rep. Prog. Theor. Phys. 29 (1966) 255.
- [15] J.R. Chaudhuri, P. Chaudhury, S. Chattopadhyay, J. Chem. Phys. 130 (2009) 234109.
- [16] A.V. Chechkin, R. Klages, J. Stat. Mech. (2009) L03002.
- [17] N.G. Van Kampen, Stochastic Processes in Physics and Chemistry, Elsevier, Amsterdam, 2006.
- [18] H. Xu, et al., Phys. Rev. Lett. 99 (2007) 204501.
- [19] G.S. Lewis, H.L. Swinney, Phys. Rev. E 59 (1999) 5457.
- [20] A.N. Kolmogorov, J. Fluid Mech. 13 (1962) 82.
- [21] D. Selmeczi, et al., Biophys. J. 89 (2005) 912.
- [22] F. Peruani, L.G. Morelli, Phys. Rev. Lett. 99 (2007) 010602.