IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Fokker–Planck description for a linear delayed Langevin equation with additive Gaussian noise

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2016 J. Phys. A: Math. Theor. 49 384002

(http://iopscience.iop.org/1751-8121/49/38/384002)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 137.222.114.243
This content was downloaded on 09/03/2017 at 12:39

Please note that terms and conditions apply.

You may also be interested in:

Magnetic Excitations and Geometric Confinement: Classical spin dynamics simulations
G M Wysin

Symmetry and Collective Fluctuations in Evolutionary Games: Symmetry and collective fluctuations: large deviations and scaling in population processes
E Smith and S Krishnamurthy

Extended Plefka expansion for stochastic dynamics
B Bravi, P Sollich and M Opper

Nonequilibrium temperature response for stochastic overdamped systems
G Falasco and M Baiesi

Diffusion properties of active particles with directional reversal
R Großmann, F Peruani and M Bär

Resilience of ${\mathscr{P}}{\mathscr{T}}$ symmetry against stochasticity in a gain-loss balanced oscillator
Mirko Lukovi, Patrick Navez, Giorgos P Tsironis et al.

Resetting of fluctuating interfaces at power-law times
Shamik Gupta and Apoorva Nagar

Expectation propagation for continuous time stochastic processes
Botond Cseke, David Schnoerr, Manfred Opper et al.

Fokker-Planck equation with linear and time dependent load forces
Kwok Sau Fa

OPEN ACCESS
IOP Publishing
Journal of Physics A: Mathematical and Theoretical
J. Phys. A: Math. Theor. **49** (2016) 384002 (29pp)
doi:10.1088/1751-8113/49/38/384002

Fokker–Planck description for a linear delayed Langevin equation with additive Gaussian noise

Luca Giuggioli[1,2,3], **Thomas John McKetterick**[1], **V M Kenkre**[4] **and Matthew Chase**[4]

[1] Bristol Centre for Complexity Sciences, University of Bristol, Bristol BS8 1UG, UK
[2] Department of Engineering Mathematics, University of Bristol, Bristol BS8 1UG, UK
[3] School of Biological Sciences, University of Bristol, Bristol BS8 1TQ, UK
[4] Physics and Astronomy Department, University of New Mexico, Albuquerque, NM, 87131, USA

E-mail: Luca.Giuggioli@bristol.ac.uk

Received 26 May 2016, revised 19 July 2016
Accepted for publication 8 August 2016
Published 30 August 2016

Abstract

We construct an equivalent probability description of linear multi-delay Langevin equations subject to additive Gaussian white noise. By exploiting the time-convolutionless transform and a time variable transformation we are able to write a Fokker–Planck equation (FPE) for the 1-time and for the 2-time probability distributions valid irrespective of the regime of stability of the Langevin equations. We solve exactly the derived FPEs and analyze the aging dynamics by studying analytically the conditional probability distribution. We discuss explicitly why the *initially* conditioned distribution is not sufficient to describe fully out a non-Markov process as both preparation and observation times have bearing on its dynamics. As our analytic procedure can also be applied to linear Langevin equations with memory kernels, we compare the non-Markov dynamics of a one-delay system with that of a generalized Langevin equation with an exponential as well as a power law memory. Application to a generalization of the Green–Kubo formula is also presented.

Keywords: delay Langevin equation, Fokker–Planck equation for delay proceeses, generalized Langevin equation, non-Markov processes, ageing

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

dynamics, generalized Green–Kubo formula, one and two-time probability distribution delayed processes

(Some figures may appear in colour only in the online journal)

1. Introduction

Markov assumptions, at the heart of the celebrated Smoluchowski's article on the equation for a diffusing particle in an external field [1], were so fruitful for the development of the modern theory of statistical physics [2–6], that, as is well-known, the equation has become a workhorse for all sorts of calculations of stochastic phenomena. Over the last century, while Markov studies and their formalism grew and found applications in physics and beyond, the same cannot be said of non-Markov analysis, which has remained in its infancy. Part of the reasons for the lack of progress is the inherent difficulty in the description of non-Markov processes, but also because the limited resolution of the observations did not present a real need for it. In recent years, however, as experimental and statistical advances have brought the ability to extract history-dependent signatures in variety of systems, from the physical [7] and engineering [8] areas to financial processes [9], there is renewed interest to develop a general theory of non-Markov processes. It is thus natural on the 100th anniversary of Marian Smoluchoski's seminal contribution to look forward and report on a formalism to study some specific non-Markov systems.

Among those systems for which the Markov assumption does not hold, we have selected linear *delayed* systems representing a classic example where the dynamics may display a remarkably different behavior compared to their Markoffian counterparts. The interplay of noise and time delays has a profound impact on many artificial and natural systems. In engineering systems, such as communication networks and sensor-actuator feedback applications, it is well known that finite transmission speeds introduce time-lags that affect process efficiency and control [10–12]. Biological systems are also affected by lagged dynamics. Delays have in fact been measured both with whole organisms, e.g. in gene regulatory networks [13, 14] and human coordination [15], and among different organisms, e.g. in collectively moving animals [16–18]. Explaining these and other observations in economical [19] and physiological systems [20] requires modeling random processes with delayed dynamics.

Langevin equations and Fokker–Planck equations (FPEs) are widely used to study the dynamics of stochastic processes. In Markov systems, there exists a well known procedure that allows one to construct an equivalent FPE description starting from a Langevin equation (see e.g. [21, 22]). In non-Markov systems such a construction generally remains an open problem. While for Markov processes knowing the system at any (arbitrary) moment in the past allows to describe the future dynamics of the system, this is not the case for non-Markov systems that may display ageing effects. Given the history dependence of non-Markov processes, to describe the dynamics, it is necessary to have full knowledge of how the system has evolved from some time in the past. The so called *bona fide* FPE representation of a non-Markov process [23, 24] thus requires to construct the conditional probability distribution at any time in the past. A proper FPE description would thus need to contain information about the initial localization as well as the the initial preparation of the state of the system.

Specific attempts to develop a procedure to write a FPE formalism from a non-Markov Langevin equation have already appeared in the past in the context of linear time non-local Langevin equations with additive Gaussian noise. The FPE construction consists of making

use of the time-convolutionless transform, a technique developed in the '70s, to convert a time non-local Langevin equation to a time local description, and thus construct an equivalent FPE [25–29]. While focus of this construction has been directed at Langevin equations where time non-locality is given by a time convolution of the stochastic variable with a memory kernel, the so-called generalized Langevin equations (GLEs), also referred to as the fractional Langevin equation [30–32], this approach has not been fully exploited in Gaussian linear delayed Langevin equation (DLE).

A well defined procedure to find FPEs for the 1- and 2-time probability distribution equivalent to the Langevin description has been developed for the case of GLEs [23, 25, 33–35]. However, in the DLE literature, such FPEs are still lacking. Of note is that more recently an FPE for the one-delay DLE described in terms of two state variables, the delayed and non-delayed one, has been proposed [36]. Although such representation has been shown to be equivalent to a random walk description [37], and generalizations to multiple delays have been put forward [38], it has an important drawback. It links the rate of change of the 1-time probability distribution of the non-delayed variable to the moments of the 2-time joint probability distribution of both the non-delayed and delayed variables (one for each delay). As a result, an exact solution for such an FPE, so far only obtained at steady state and for the case of one delay [39, 40], cannot be found without supplementing information about the moments of the distribution from the Langevin equation. This under-determined nature of the FPE [40] calls into question its practical utility and has recently brought statements about finding the exact time dependent probability distribution equivalent to the DLE as a problem beyond reach [41].

Here we present a solution to this problem by building on recent studies of some of the present authors [42, 43]. We construct and solve the 2-time FPE including the case where the initial history is non-zero (zero initial history makes a DLE a special case of a GLE). In other words we construct and solve closed FPEs for the 1- and 2-time joint probability distributions associated with a DLE. From that we are able to write the *bona fide* FPE representation of a DLE in terms of the the conditional probability distribution of the delay process and show how ageing (see e.g. [44]) effects appear evident from the dependence of the dynamics on the initial preparation of a system. Furthermore, our exact analytics allow us to show that DLEs and GLEs do not satisfy the Markov condition and are time-inhomogeneous processes except when they reduce to the Wiener or Ornstein–Uhlenbeck process. As an application we use the formalism of the 2-time probability distribution to highlight interesting differences when computing multi-time averages of the one-delay DLE versus a GLE with exponential memory or with a power law memory. Use of the conditional probability distribution also allows us to use our formalism to construct analytically the generalization of the Green–Kubo relation.

The paper is laid out as follows. We start in section 2 by reviewing the constitutive links between joint and conditional distributions of Gaussian processes and the relations between the Fokker-Planck and Langevin description. In section 3 we construct the 1- and 2-time FPE from a DLE with any number of delays and arbitrary history function, solve them analytically and present the *bona fide* representation for the conditional probability distribution. In section 4 we show that DLE and GLE can be analyzed together given their formal similarities and that their time non-locality describe non-Markov processes. The role of time homogeneity in these non-Markov systems and their relation to the Chapman–Kolmogorov equation (CKE) is also elucidated. In section 5 we study three specific non-Markov processes, a DLE with a single delay and a GLE with exponential and power law memory. We compare them in computing conditional probabilities when the system preparation time does not coincide with the start of the observation as well in calculating multi-time averages. Analysis of the

generalized Green–Kubo formalism is also presented. Finally section 7 forms the concluding discussion.

2. Interrelation between Langevin and Fokker–Planck description

We review here some preliminary notions and the notations necessary to follow our derivation starting from the definition of the complete description of a stochastic variable $x(t)$ in terms of the so-called n-time probability distribution $W_n(x_1, t^{(1)}; \ldots; x_n, t^{(n)})$. This distribution represents the joint probability for the random variable to attain the value x_1 at time $t^{(1)}$, x_2 at time $t^{(2)}$ through to x_n at time $t^{(n)}$ with $t^{(1)} > t^{(2)} > \ldots > t^{(n)}$, while it reduces to $\delta(x_p - x_q) W_{n-1}(x_1, t^{(1)}; \ldots; x_{n-1}, t^{(n-1)})$ whenever $t^{(p)} = t^{(q)}$ where $\delta(z)$ is a Dirac delta function. The n-time probability distribution allows one to compute any m-time average with $m \leqslant n$ [45]. While the $(n-1)$-time probability distribution $W_{n-1}(x_1, t^{(1)}; \ldots; x_{n-1}, t^{(n-1)})$ is simply the marginal of the n-time distribution, the conditional distribution $Q_{n|m}$ of the variables x_{m+1} at time $t^{(m+1)}$ and x_{m+n} at times $t^{(m+n)}$, given the variables x_1 at time $t^{(1)}$ and x_m at time $t^{(m)}$, is given by Bayes theorem [46] as the ratio of the joint m- and $(m+n)$-time distributions (with $n, m \geqslant 1$) or alternatively the ratio of the conditional r- and $(n+r)$-time distributions (with $r \geqslant 1$):

$$Q_{n|m}(x_{m+1}, t^{(m+1)}; ..; x_{m+n}, t^{(m+n)}|x_1, t^{(1)}; ..; x_m, t^{(m)}) = \frac{W_{n+m}(x_1, t^{(1)}; ..; x_{m+n}, t^{(m+n)})}{W_m(x_1, t^{(1)}; ..; x_m, t^{(m)})}$$
$$= \frac{Q_{n+r|m-r}(x_{m-r+1}, t^{(m-r+1)}; ..; x_{m+n}, t^{(m+n)}|x_1, t^{(1)}; ..; x_{m-r}, t^{(m-r)})}{Q_{r|m-r}(x_{m-r+1}, t^{(m-r+1)}; ..; x_m, t^{(m)}|x_1, t^{(1)}; ..; x_{m-r}, t^{(m-r)})}. \tag{1}$$

The general relation between the joint n-time probability $W_n(x_1, t^{(1)}; \ldots; x_n, t^{(n)})$ and $P_n(x_1, t^{(1)}; \ldots; x_n, t^{(n)}|x_0, 0) = Q_{n|1}(x_1, t^{(1)}; \ldots; x_n, t^{(n)}|x_0, 0)$, the *initially* conditioned probability, is found via marginalisation of Bayes relation (1). From $Q_{n|1} = W_{n+1}/W_1$, when W_1 is restricted to the initial distribution $W_0(z) = W_1(z, 0)$, one has $W_n(x_1, t^{(1)}, \ldots, x_n, t^{(n)}) = \int \mathrm{d}z\, P_n(x_1, t^{(1)}, \ldots, x_n, t^{(n)}|z, 0) W_0(z)$, which indicates the exact equivalence between W_n and P_n when the system is prepared with a localized initial distribution $W_0(z) = \delta(z - x_0)$.

The unique class of stochastic processes called Markov are those for which the conditional probability depends only on the most recent previous times [27, 47, 48]. This means that they satisfy the relation

$$Q_{n|m}(x_{m+1}, t^{(m+1)}; ..; x_{m+n}, t^{(m+n)}|x_1, t^{(1)}; ..; x_m, t^{(m)})$$
$$= Q_{n|1}(x_{m+1}, t^{(m+1)}; ..; x_{m+n}, t^{(m+n)}|x_m, t^{(m)}), \tag{2}$$

for $t^{(1)} < t^{(m)} < t^{(m+n)}$. With Bayes theorem the Markov property (2) then implies that $W_n = Q_{1|1} W_{n-1}$ for $n \geqslant 1$. Recursive use of this identity shows that W_n can be expressed only in terms of $Q_{1|1}$ and W_1. Since both W_1 and $Q_{1|1}$ can be obtained from W_2, i.e. $W_1 = \int W_2$ and $W_2 = Q_{1|1} W_1$, Markov processes are completely characterized by W_2.

Markov processes are time-homogeneous if they have shift invariant conditional probabilities, i.e. $Q_{1|1}(x, t+s|x', s) = Q_{1|1}(x, t|x', 0)$. As Markov processes are completely determined by $Q_{1|1}$ and W_1, it follows that time-homogeneous Markov processes are completely determined by the initially conditioned probability $P_1(x, t|x_0, 0) = Q_{1|1}(x, t|x' = x_0, t' = 0)$. In contrast, non-Markov processes are in general only completely described by the full hierarchy of joint probability distributions. However, Gaussian processes represent an exception to this rule as all n-time averages of a Gaussian distribution can be

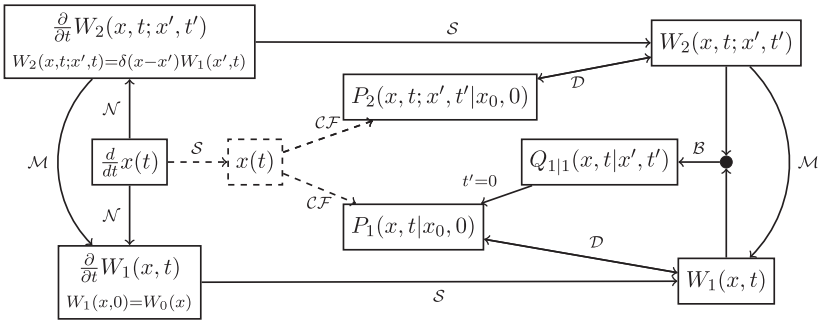

Figure 1. From a time non-local linear Gaussian Langevin equation for a stochastic variable $x(t)$, it is possible to derive, by solving ($\mathcal{S}$) the equation, performing a time-convolutionless transform (see section 3) and using Novikov's theorem ($\mathcal{N}$), Fokker–Planck equations for both the joint probabilities $W_1(x, t)$ and $W_2(x, t; x', t')$. The symbol $\mathcal{M}$ indicates that the two objects are linked by marginalization. Solving these equations ($\mathcal{S}$) and applying Bayes theorem ($\mathcal{B}$) gives the conditional probability, $Q_{1|1}(x, t|x', t')$, which gives, in the limit $t' = 0$, the initially conditioned probability, $P_1(x, t|x', 0)$. Alternatively $P_1(x, t|x', 0)$ and $P_2(x, t; x', t'|x'', 0)$ can be obtained directly from the solution of the Langevin equation using characteristic functionals ($\mathcal{CF}$) (see e.g. [42]). The double-arrowed line ($\mathcal{D}$) aims to indicate the relations $W_1(x, t) = \int \mathrm{d}z P_1(x, t|z, 0) W_0(z)$ and $W_2(x, t; x', t') = \int \mathrm{d}z\, P_2(x, t, x', t'|z, 0) W_0(z)$ between the initially conditioned and the joint probability distributions.

found from the 1- and 2-time averages [49]. Consequently, Gaussian non-Markov processes, like Markov processes, are completely determined by the 2-time joint probability distribution, $W_2(x, t; x', t')$. In figure 1 we display, as a visual aid, the general interrelations between $W_2(x, t; x', t')$, $W_1(x, t)$, $Q_{1|1}(x, t|x', t')$, $P_1(x, t|x_0, 0)$ and $P_2(x, t; x', t'|x_0, 0)$ for linear Gaussian processes. In this article we focus primarily on developing a formalism of what those arrows represent when the Langevin equation $\frac{\mathrm{d}x(t)}{\mathrm{d}t}$ is a DLE, but we also show that the formalism can be applied in general to GLEs. We construct and solve the associated FPE, but we do not report on the method of characteristic functionals (dashed lines), which has already been presented in [42, 50].

3. DLE and a *bona fide* Fokker–Planck representation

3.1. From time non-local to time local Langevin equations

A general DLE, for simplicity considered 1D here, can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -\sum_{i=1}^{N}\gamma_i x(t-\tau_i) + \sigma\eta(t). \tag{3}$$

In equation (3) $\eta(t)$ represents Gaussian white noise of strength σ, i.e. $\langle\eta(t)\rangle = 0$ and $\langle\eta(t)\eta(t')\rangle = \delta(t-t')$, while the coefficients γ_i are rates of decay for each of the N-delays, $0 < \tau_1 < \ldots < \tau_N$, and a given history function $x(t) = \beta(t)$ for $-\tau_N \leqslant t < 0$.

The deterministic dynamics of equation (3) is captured formally by the Green's functions $\lambda(t)$ satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda(t) = -\sum_{i=1}^{N}\gamma_i\lambda(t-\tau_i), \tag{4}$$

with $\lambda(0) = 1$ and $\lambda(t) = 0$ for all $t < 0$. The dynamics of $x(t)$ can be written explicitly in terms of this Green's function as [51]

$$x(t) = x_0\lambda(t) + \sigma\int_0^t \mathrm{d}s\ \lambda(t-s)\eta(s) + \Psi_d(t), \tag{5}$$

where the first two terms are present in the solutions of (3) with a dependence on the history function given by

$$\Psi_d(t) = -\sum_{i=1}^{N}\gamma_i\int_{-\tau_i}^{0}\mathrm{d}s\ \lambda(t-\tau_i-s)\beta(s), \tag{6}$$

with $\Psi_d(0) = 0$.

The time-convolutionless transform consists of converting the Langevin equation into a time-local form by differentiating the solution (5) with respect to time and substituting x_0 from equation (5) into the resulting equation [25–27, 34]. The DLE then reduces to the time local form

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -A(t)x(t) + B_d(t) + \sigma\xi(t), \tag{7}$$

with $x(0) = x_0$ and

$$A(t) = -\frac{\dot{\lambda}(t)}{\lambda(t)}, \tag{8}$$

$$B_d(t) = -\frac{\dot{\lambda}(t)}{\lambda(t)}\Psi_d(t) + \dot{\Psi}_d(t), \tag{9}$$

$$\xi(t) = \lambda(t)\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{d}s\ \frac{\lambda(t-s)}{\lambda(t)}\eta(s). \tag{10}$$

The non-Markoffian characteristics of the equation, which was apparent in the original DLE, is now lumped in the time-dependent terms $A(t)$ and $B_d(t)$ and the colored noise obeying

$$\langle\xi(t)\rangle = 0,$$

$$\langle\xi(t)\xi(t')\rangle_{t'<t} = \lambda(t)\lambda(t')\frac{\mathrm{d}}{\mathrm{d}t'}\left(\frac{1}{\lambda(t')}\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{\lambda(t)}\int_0^{t'}\mathrm{d}s\ \lambda(t-s)\lambda(t'-s)\right]\right). \tag{11}$$

3.2. FPE for the 1 and 2-time probability distributions

The procedure to construct a FP equation for the 1- and 2-time probability distributions from equations (7)–(11) consists of expressing the dynamics of $W_2(x, t; x', t') = \langle\delta(x - x(t))\delta(x' - x(t'))\rangle$ and $W_1(x, t) = \langle\delta(x - x(t))\rangle$ as (see e.g. [21])

$$\frac{\partial}{\partial t}W_1(x, t) = -\frac{\partial}{\partial x}\left\langle\frac{\mathrm{d}x(t)}{\mathrm{d}t}\bigg|_{x(t)=x}\delta(x - x(t))\right\rangle, \tag{12}$$

$$\frac{\partial}{\partial t}W_2(x,\,t;\,x',\,t') = -\frac{\partial}{\partial x}\left\langle \frac{\mathrm{d}x(t)}{\mathrm{d}t}\bigg|_{x(t)=x} \delta(x - x(t))\delta(x' - x(t'))\right\rangle. \quad (13)$$

In the Markoffian case one would evaluate the required ensemble average by substituting equation (3) into the above expressions and using properties of the Dirac delta to exchange the derivative of the stochastic variable $x(t)$ for that of the fixed variable x. However the time non-local form of equation (3) implies that this substitution cannot be used to evaluate the required averages [43]. Performing this procedure leads to a non-closed FPE [52], that is a partial differential equation where the first time derivative of $W_1(x,\,t)$ is dependent on moments of $W_2(x,\,t;\,x',\,t')$. Although in appendix A we show that it is possible to find a solution of such a FPE by postulating an ansatz informed by the dynamics of the moments of the Langevin equation, here we construct a closed form FPE.

To construct and solve as an initial value problem a *bona fide* FPE we need to use instead the time-local Langevin equation (7). Its substitution into equations (12) and (13) allows the evaluation of the averages to obtain the drift and diffusion terms of the FPEs. While the drift term is readily found to be equivalent to the deterministic term in the Langevin equation (with a change of sign), finding the diffusion term requires calculating the functional derivative of $x(t)$ with respect to $\xi(t)$ with the help of Novikov's theorem [53] (see also [54, 55]). For the particular case of noise obeying fluctuation–dissipation the details of this procedure are given in [29] and the drift and diffusion terms of the 1-time FPE are shown to be equal. For our case of delta-correlated noise these terms are found to be very different. In appendix B we explicitly show their derivation for both the 1-time (also shown in [34]) and the more complicated 2-time distributions. The resulting FPEs are

$$\frac{\partial}{\partial t}W_1(x,\,t) = \frac{\partial}{\partial x}\left[A(t)x - B_d(t) + D(t)\frac{\partial}{\partial x}\right]W_1(x,\,t), \quad (14)$$

$$\frac{\partial}{\partial t}W_2(x,\,t;\,x',\,t') = \frac{\partial}{\partial x}\left[A(t)x - B_d(t) + C(t',\,t)\frac{\partial}{\partial x'} + D(t)\frac{\partial}{\partial x}\right]W_2(x,\,t;\,x',\,t'), \quad (15)$$

where

$$D(t) = \frac{\sigma^2\lambda^2(t)}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{d}s\,\frac{\lambda^2(s)}{\lambda^2(t)}, \quad (16)$$

$$C(t',\,t) = \sigma^2\lambda(t)\frac{\mathrm{d}}{\mathrm{d}t}\int_0^{t'} \mathrm{d}s\frac{\lambda(t-s)\lambda(t'-s)}{\lambda(t)}. \quad (17)$$

When $t' = 0$ the joint distribution reduces to the 1-time distribution and the two FPEs are identical since $C(t' \to 0,\,t) \to 0$.

Compared to the usual FPE equations (14) and (15) display explicit time-dependence in the diffusion constant $D(t)$, in the strength of the parabolic potential $A(t)$ and in the drift terms both along x and x'. Moreover, if the potential and the diffusion constant changes sign, the FPEs may describe diffusive $D(t) > 0$ and anti-diffusive $D(t) < 0$ dynamics within a parabolic potential that may vary from being attractive $A(t) > 0$ to being repulsive $A(t) < 0$. Moreover, in the parameter regime for which the delay Langevin Green's function display oscillatory behavior, the magnitude of the FPE coefficients may actually become unbounded. Hänggi [47, 56] pointed out that this potential scenario may occur when representing non-Markov processes. The origin of this divergence is due to the convolutionless transform not being defined whenever $\lambda(t) = 0$. This has the consequence of having the $A(t)$, $B_d(t)$, $D(t)$

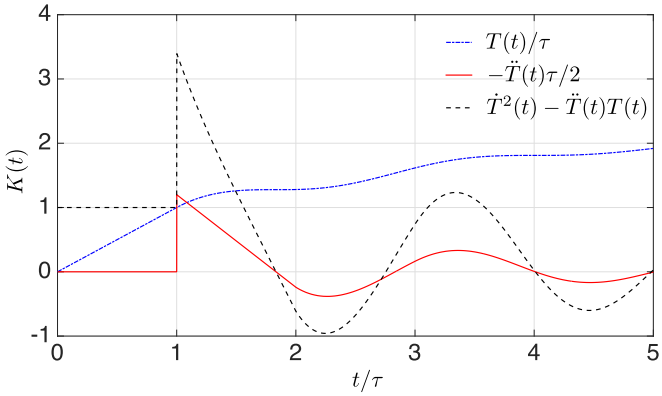

Figure 2. Time dependence of the rescaled potential strength $A(t)$ and diffusion coefficient $D(t)$ in the FPE (14) and (15) for the case in which the delay Langevin equation (3) with the single delay τ displays oscillatory dynamics ($\gamma\tau = 1.2$). The label $K(t)$ indicates either of the dimensionless quantities $T(t)/\tau$, $-\ddot{T}(t)\tau/2$ or $\dot{T}^2(t) - \ddot{T}(t)T(t)$. As the zeros of $\lambda(t)$ give rise to flex points in the function $T(t)$, the change of sign from positive to negative of the function $\ddot{T}(t)$ and $\dot{T}^2(t) - \ddot{T}(t)T(t)$ coincide, whereas they do not when crossing from negative to positive since the maxima of $T(t)$ are not flex points. The former case corresponds to the time values when $A(t)$ and $D(t)$ diverge simultaneously, while the latter case indicates the separate crossings of the zero value for $A(t)$ and $D(t)$.

and $C(t', t)$ terms becoming infinitely large at those specific moments in time. However, the manner in which the sign of the coefficients become unbounded is such that the probability distributions are well behaved, e.g. the coefficient $A(t)$ grows unbounded with the same sign as $D(t)$. To display the balancing effects of the sign change in the FPE coefficients we display in figure 2 the rescaled coefficients

$$T(t) = \int_0^t \mathrm{d}s\ \lambda^2(s), \tag{18}$$

$$\ddot{T}(t) = 2\lambda(t)\dot{\lambda}(t) = 2\lambda^2(t)A(t), \tag{19}$$

$$\dot{T}^2(t) - \ddot{T}(t)T(t) = \frac{2\lambda^2(t)D(t)}{\sigma^2}. \tag{20}$$

To obtain a FPE description with bounded coefficients also in those cases when the dynamics of $\lambda(t)$ is oscillatory, we make a variable transformation from t to the rescaled time $T(t)$ and rewrite equations (14) and (15) in terms of $\overline{W}_1(x, T) = W_1(x, T(t))$ and $\overline{W}_2(x, T; x', T') = W_2(x, T(t); x', T(t'))$ as

$$\frac{\partial}{\partial T}\overline{W}_2(x, T; x', T') = \frac{\partial}{\partial x}\left\{\overline{A}(T)x - \overline{B}_d(T) + \overline{C}(T', T)\frac{\partial}{\partial x'} + \overline{D}(T)\frac{\partial}{\partial x}\right\}\overline{W}_2(x, T; x', T'), \tag{21}$$

$$\frac{\partial}{\partial T}\overline{W}_1(x,\,T) = \frac{\partial}{\partial x}\left[\overline{A}\,(T)x - \overline{B}_d(T) + \overline{D}\,(T)\frac{\partial}{\partial x}\right]\overline{W}_1(x,\,T), \tag{22}$$

where

$$\begin{aligned}
\overline{A}\,(T) &= -\,\lambda(t^{-1}(T))\,\dot{\lambda}(t^{-1}(T))\\
\overline{C}\,(T',\,T) &= \lambda^2(t^{-1}(T))C\,(t'^{-1}(T'),\,t^{-1}(T)),\\
\overline{B}_d(T) &= -\,\lambda(t^{-1})\,\dot{\lambda}(t^{-1}(T))\Psi_d(t^{-1}(T)) + \lambda^2(t^{-1}(T))\dot{\Psi}_d(t^{-1}(T))\\
\overline{D}\,(T) &= \lambda^2(t^{-1}(T))D\,(t^{-1}(T))
\end{aligned} \tag{23}$$

with $t^{-1}(T)$ and $t'^{-1}(T')$ being respectively the inverse functions of $T(t)$ and $T\,(t')$. Equations (21) and (22) supplemented, respectively, by the initial preparation of the system $\overline{W}_1(x,\,0) = W_0(x)$ and by the initial condition $\overline{W}_2(x,\,T';\,x',\,T') = \delta\,(x - x')\overline{W}_1(x',\,T')$ constitute the sought after *bona fide* representation of the delay Langevin equation (3) in the rescaled time T and T'. The dynamics in the actual time t is finally obtained by re-expressing the time T as $T(t)$. This time transformation amounts to a rewriting of the Langevin equation where the dynamics is described in time T rather than time t.

To gain further insight we make a variable transformation (see appendix C) and solve exactly equations (14) and (15) and write their analytic solutions in time t and t' as

$$W_1(x,\,t) = [2\pi\nu(t)]^{-1/2}\int_{-\infty}^{\infty}\mathrm{d}z\;W_0(z)\exp\left\{-\frac{(x - z\lambda(t) - \Psi_d(t))^2}{2\nu(t)}\right\}, \tag{24}$$

$$\begin{aligned}
W_2(x,\,t;\,x',\,t') &= \frac{1}{\sqrt{4\pi^2\nu(t')\nu(t)[1 - r^2(t',\,t)]}}\int_{-\infty}^{\infty}\mathrm{d}z\;W_0(z)\\
&\times\exp\left\{\frac{-1}{2[1 - r^2(t',\,t)]}\left[\left(\frac{x - z\lambda(t) - \Psi_d(t)}{\nu^{1/2}(t)}\right)^2 + \left(\frac{x' - z\lambda(t') - \Psi_d(t')}{\nu^{1/2}(t')}\right)^2\right.\right.\\
&\left.\left. - 2r(t',\,t)\left(\frac{x - z\lambda(t) - \Psi_d(t)}{\nu^{1/2}(t)}\right)\left(\frac{x' - z\lambda(t') - \Psi_d(t')}{\nu^{1/2}(t')}\right)\right]\right\},
\end{aligned} \tag{25}$$

where $t \geqslant t'$ and

$$\nu(t) = \sigma^2 T\,(t), \qquad r(t',\,t) = \frac{\sigma^2\int_0^{t'}\mathrm{d}s\;\lambda(t - s)\lambda(t' - s)}{\sqrt{\nu(t)\nu(t')}}. \tag{26}$$

Equation (25) is a bivariate Gaussian with mean $z\lambda(t) + \Psi_d(t)$, mean square displacement (MSD) or variance $\nu(t)$, 2-time correlation function $r(t',\,t)$ and covariance $r(t',\,t)\sqrt{\nu(t)\nu(t')}$. The 2-time correlation function $r(t',\,t)$ is defined for $t \geqslant t'$ with $r(0,\,t) = 0$ and $r(t',\,t \to +\infty) = 0$ and equals 1 whenever $t = t'$. For the special case in which all $\tau_i = 0$, the Langevin equation (4) describes the Ornstein–Uhlenbeck process and the correlation function is simply the ratio of the function $h(t')/h(t)$ where $h(t) = \sqrt{\left[\exp\left(2\sum_i\gamma_i\,t\right) - 1\right]}$, whereas for Wiener processes $h(t) = \sqrt{t}$. A rescaling of t and t' in (24) and (25) to T and T' would give the corresponding analytic form of $\overline{W}_1$ and $\overline{W}_2$.

3.3. Conditional probability and aging dynamics

The 1- and 2-time probability distributions (24) and (25) represent the time-dependent equivalent description of the DLE in equation (3) and they allow to derive the conditional

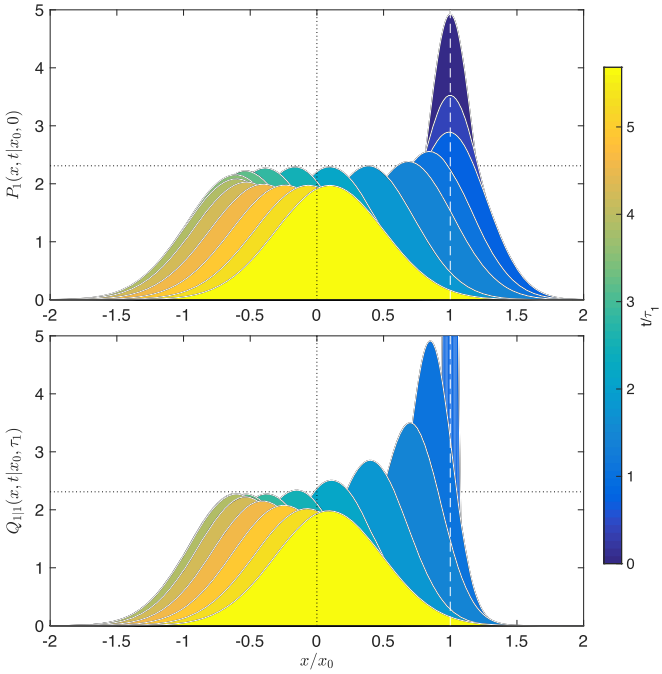

Figure 3. Effect of the ageing dynamics on the probability distribution of the 2-delay linear Gaussian process governed by equation (27) with the Green's function (28) for $\gamma\tau_1 = 1/2$ and $\gamma\tau_2 = 3/4$ and with history function $\beta(t) = 0$ ($B_d(t) = 0$). In the top and bottom panels we plot at different times equation (24) with a Dirac delta initial condition centered at $x = x0$, that is the dynamics of the initially conditioned probability distribution $Q_{1|1}(x, t|x', 0)$. In the bottom panel we plot $Q_{1|1}(x, t|x', t')$ from equation (27) with the same initial preparation as in the top panel, starting at $t' = \tau_1$ and $x' = x_0$ (the units of the vertical axes are arbitrary). The snapshots in both panels are taken at equally spaced intervals of $\tau_1/3$ starting at $t = \tau_1/3 + \tau_2/100$ in the top panel and $t = \tau_1 + \tau_2/100$ in the bottom panel. The vertical dashed line at $x = x_0$ indicates where the system was initially localized, whereas the horizontal dotted line drawn at a value of 2.31 is displayed to help visualize the difference in the broadening and drift dynamics of $Q_{1|1}(x, t|x', 0)$ and $Q_{1|1}(x, t|x_0, \tau_1)$ between the time $t/\tau_1 = 121/60$ and $t/\tau_1 = 201/60$ in the top (from the 6th to the 10th snapshot) and bottom panel (from the 4th to the 8th snapshot).

probability $Q_{1|1}(x, t|x', t') = W_2(x, t; x', t')/W_1(x', t')$ given by

$$
\begin{aligned}
Q_{1|1}(x, t|x', t') = &\left[\sqrt{\nu(t)[1 - r^2(t', t)]}\int_{-\infty}^{\infty} \mathrm{d}y\, W_0(y)\right.\\
&\left.\times \exp\left\{-\frac{[x' - y\lambda(t') - \Psi_d(t')]^2}{2\nu(t')}\right\}\right]^{-1} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}z\, W_0(z)\\
&\times \exp\left\{-\frac{[x - z\lambda(t) - \Psi_d(t)]^2}{2\nu(t)[1 - r^2(t', t)]} - \frac{[x' - z\lambda(t') - \Psi_d(t')]^2}{2\nu(t')[1 - r^2(t', t)]}\right.\\
&\left.+ \frac{r(t', t)[x - z\lambda(t) - \Psi_d(t)][x' - z\lambda(t') - \Psi_d(t')]}{\nu^{1/2}(t)\nu^{1/2}(t')[1 - r^2(t', t)]}\right\}. \qquad (27)
\end{aligned}
$$

In equation (27) the dependence of $Q_{1|1}(x, t|x', t')$ on the initial preparation of the system appears explicitly in the convolution integral with $W_0(z)$ both in the numerator and denominator. When the initial preparation affects the dynamics between time t' and t, the system is said to undergo aging dynamics, a situation that does not occur in Markov cases. Indeed one can show that for an Ornstein–Uhlenbeck or a Wiener process, that is when $\Psi_d(t) = 0$ and $\lambda(t) = \mathrm{e}^{-\mathcal{A}t}$ for $\mathcal{A} \geqslant 0$, the initial preparation dependence is lost since $Q_{1|1}(x, t|x', t') = [2\pi\sigma^2(1 - \lambda^2(t)/\lambda^2(t'))/\mathcal{A}]^{-1} \exp\{-[x - x'\lambda(t)/\lambda(t')]^2/[2\sigma^2(1 - \lambda^2(t)/\lambda^2(t'))/\mathcal{A}]\}$. To display aging effects we consider the non-Markov scenario in which the system dynamics is governed by the Langevin equation (3) with two delays τ_1 and τ_2 and, for simplicity, one single rate $\gamma = \gamma_1 = \gamma_2$. We use the general analytic expression of the Green's function for any number of delays [42] and write

$$\lambda(t) = \sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty} \frac{[-\gamma(t - k_1\tau_1 - k_2\tau_2)]^{k_1+k_2}}{k_1!k_2!}\Theta(t - k_1\tau_1 - k_2\tau_2), \tag{28}$$

where $\Theta(z)$ is the Heaviside step function. Equation (28), despite its complicated appearance, is not composed of infinite series since the summations get truncated whenever the argument of the Heaviside function is negative, and is well suited to evaluate $\lambda(t)$ at short and intermediate times. Compared to the Green's function for the case with one delay [50, 57, 58], for which the Green's function is a polynomial of degree k_1 within the interval $k_1\tau_1 \leqslant t < (k_1 + 1)\tau_1$, for higher number of delays the degree of the polynomial of $\lambda(t)$ depends on the combination of k_j values. Its order may in fact increase or remain unchanged as time increases. To detect clearly the differences due to aging in our delay Langevin equation we compare the dynamics governed by the two delay Green's function (28) in the oscillatory regime, which is generally present in parameter space close to where the system becomes unstable. As the system exhibits an unstable regime for values of $\gamma\tau_1$ and $\gamma\tau_2$ beyond the parameter space boundary function $\cos[(\pi/2)(\tau_1 - \tau_2)/(\tau_1 + \tau_2)] = \pi/[2\gamma(\tau_1 + \tau_2)]$ [42], it turns out that with the choice of $\gamma\tau_1 = 1/2$ and $\gamma\tau_2 = 3/4$ close to the unstable regime we ensure that the system is in the oscillatory regime. As the mean position in equation (27) is largely controlled by $\lambda(t)$, when $\lambda(t)$ changes sign the probability distributions reverse the drift direction. Since $T(t) = 0$ either when $\lambda(t) = 0$ or when $\mathrm{d}\lambda(t)/\mathrm{d}t = 0$, as time evolves the broadening of the probability distribution can be quite different depending on the type of zero crossing of $T(t)$. When approaching the former case (inflection points), the change of width slows down, whereas it changes more rapidly as time approaches those values for which $\mathrm{d}\lambda(t)/\mathrm{d}t = 0$. We thus have that $Q_{1,1}(x, t|x_0, 0)$ and $Q_{1|1}(x, t|x, t')$ may decelerate or accelerate the broadening of their width when close to the time at which, respectively, $\lambda(t) = 0$ or $\mathrm{d}/\mathrm{d}t\lambda(t) = 0$. For our 2-delay system this is clearly visible in the top panel of figure 3 where $Q_{1,1}(x, t|x_0, 0)$ is plotted for equally spaced time intervals (snapshots). Since $\lambda(t) = 1$ for times shorter than the smallest delay, the distribution broadens without shifting in the first three snapshots, while it also starts moving to the left in the 4th and 5th snapshots. The height of the Gaussian gets only mildly reduced in the sequence of snapshots from the 6th to the 10th indicating a slowing down of the broadening which occurs before and after time $t/\tau_1 \simeq 2.5$, which is when $\lambda(t) = 0$. As time approaches $t/\tau_1 \simeq 3.82$ (between the 11th and the 12th snapshot) the rate of change of the width of the Gaussian is highest. It corresponds to when $\mathrm{d}\lambda(t)/\mathrm{d}t = 0$ and reversal of the direction of movement occurs. Spreading of $Q_{1,1}(x, t|x_0, 0)$ slows down again in the last four snapshots, in correspondence to when the mean of the Gaussian crosses the origin ($t/\tau_1 \simeq 5.5$), and before the second direction reversal (not shown) at $t/\tau_1 \simeq 6.83$.

This direction reversal and slowing down/speeding up of the Gaussian width is also observed with $Q_{1|1}(x, t|x_0, \tau_1)$, but the dynamics differ considerably since the MSD is now given by $\nu(t)[1 - r^2(\tau_1, t)]$. The distribution drifts to the left from the start and its width increases without evident signs of slowing down except when close to reversing direction. This is nearly opposite to what occurs with $Q_{1,1}(x, t|x_0, 0)$ when the rate of spread is at its maximum upon reversing direction. The last snapshot on the other hand indicates that the two distributions are becoming similar as we expect since $r(\tau_1, t) \to 0$ for large t.

4. Non-Markov processes and time inhomogeneity

Armed with exact expressions for the conditional probability, we can find the conditions under which the Langevin equation describe Markov processes by comparing $Q_{1|1}(x, t|x't')$ with $Q_{1|2}(x, t|x', t'; x'', t'')$. To avoid constructing $W_3(x, t; x', t'; x'', t'')$ to derive the most general conditional distribution $Q_{1|2} = W_3/W_2$, we calculate the specific one in which one of the conditioned time is the initial time. From (1) we know that $Q_{1|2}(x, t|x', t'; x_0, 0) = Q_{2|1}(x, t; x', t'|x_0, 0)/Q_{1|1}(x, t|x_0, 0)$ and that $Q_{2|1}(x, t; x', t'|x_0, 0) = P_2(x, t; x', t'|x_0, 0)$ with the latter being identical with $W_2(x, t; x', t')$ when the initial preparation $W_0(z) = \delta(x - x_0)$. We can thus write

$$Q_{1|2}(x, t|x', t'; x_0, 0) = \frac{\exp\left\{-\frac{\left(x - x_0\lambda(t) - \Psi_d(t) - r(t', t)\frac{\nu^{1/2}(t)}{\nu^{1/2}(t')}[x' - x_0\lambda(t') - \Psi_d(t')]\right)^2}{2[1 - r^2(t', t)]\nu(t)}\right\}}{\sqrt{2\pi\nu(t)[1 - r^2(t', t)]}}. \tag{29}$$

Obviously $Q_{1|1}(x, t|x', t')$ and $Q_{1|2}(x, t|x', t'; x_0, 0)$ are not equivalent (for $t' > 0$) and thus do not satisfy the Markov property (2). The non-Markoffian character of the process can be evinced by the explicit dependence on x_0 in equation (29), which disappears only if $\frac{r(t', t)\nu^{1/2}(t)}{\nu^{1/2}(t')} = \frac{\lambda(t)}{\lambda(t')}$ or, equivalently, if $\lambda(t')\int_0^t \mathrm{d}p\ \lambda(p)\lambda(p + t' - t) = \lambda(t)\int_0^{t'} \mathrm{d}s\ \lambda^2(s)$. By differentiating with respect to t the latter expression and using the fact that $\int_0^t \mathrm{d}p\ \lambda(p)\mathrm{d}/\mathrm{d}t[\lambda(p + t' - t)] = -\lambda(t)\mathrm{d}/\mathrm{d}t'\left[\int_0^{t'} \mathrm{d}s\ \lambda^2(s)/\lambda(t')\right]$, one obtains the condition $[\mathrm{d}/\mathrm{d}t'\lambda(t')]/\lambda(t') = [\mathrm{d}/\mathrm{d}t\lambda(t)]/\lambda(t)$. This relation can be satisfied only when $\lambda(t) = \exp(-\mathcal{A}\,t)$ for $\mathcal{A} \geqslant 0$. With this time dependence for $\lambda(t)$ equation (29) loses the dependence on $\Psi_d(t)$ and $\Psi_d(t')$, in addition to its x_0-dependence. We can thus state that $Q_{1|1}(x, t|x', t') = Q_{1|2}(x, t|x', t'; x_0, 0)$ only when the Green's function of the Langevin equation is of the form $\exp(-\mathcal{A}\,t)$ describing a Wiener ($\mathcal{A} = 0$) or a Ornstein–Uhlenbeck process ($\mathcal{A} > 0$).

With this understanding we can also ask whether the processes described by equation (27) are time-homogeneous. Inspection of equation (27) for $W_0(x) = \delta(x - x')$ shows that $Q_{1|1}(x, t|x', t') = Q_{1|1}(x, t - t'|x', 0)$ only if $\lambda(t - s)\lambda(s) = \lambda(t)$. By Laplace transformation one can show that this identity is satisfied once again only if $\lambda(t) = \exp(-\mathcal{A}\,t)$ for any $\mathcal{A} \geqslant 0$. In other words, we conclude that conditional probability (27) represents time-inhomogeneous non-Markov processes, which reduce to the time homogeneous Ornstein–Uhlenbeck or Wiener processes when time non-locality in the Langevin equation is eliminated.

With our analytic results we can determine whether the conditional probability satisfies the CKE [59, 60], i.e. whether $Q_{1|1}(x, t|z, t'') = \int \mathrm{d}x' Q_{1|1}(x, t|x', t')Q_{1|1}(x', t'|x'', t'')$. While it is known that satisfying CKE is a necessary but not a sufficient condition of all Markov processes [48, 61], non-Markov processes may or may not obey it [47]. Direct substitution of

the expression (27) shows that $Q_{1|1}(x, t|x', t')$ satisfies CKE only if the 2-time correlation function is subject to the requirement $r(t'', t) = r(t', t)r(t'', t')$ with $t \geqslant t' \geqslant t''$. This single criterium is valid for any $Q_{1|1}$ derived from a 2-time bivariate Gaussian and agrees with the three simultaneous conditions derived by Hanggi for a general Gaussian process to satisfy CKE [56].

5. GLEs and multi-time average analysis

Although we have derived a FPE by starting from a DLE, the same technical steps need to be followed to construct the *bona fide* representation of the GLE

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -\gamma \int_0^t \mathrm{d}t'\, \phi(t - t')x(t') + \sigma\eta(t). \tag{30}$$

The local Langevin equation constructed from a GLE would be identical to equation (7) except for the absence of the $B_d(t)$ term which carries the information of the history function $\beta(t)$. With $\Psi_d(t) = 0$ equations (24) and (25) have in fact the same formal appearance of the 1- and 2-time probability distribution that one would derive starting from a GLE. The only difference is in the functional dependence of the Green's function $\lambda(t)$, which satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\lambda(t) = -\gamma \int_0^t \mathrm{d}t'\, \phi(t - t')\lambda(t'). \tag{31}$$

Indeed a GLE reduces to a DLE when $\phi(t) = \delta(t - \tau)$ or more generally to the multiple delay differential equation (4) when $\phi(t) = \sum_{i=1}^{N} \gamma_i \delta(t - \tau_i)$ but only if the history function $\beta(t)$ is identically zero [43].

In section 3.3 we have shown the qualitative difference in the dynamics of a stochastic linear delay system when the preparation and observation times do not coincide. Here we highlight the time-inhomogeneous nature of the (non-Markov) DLE and GLE, even when preparation and initial observation times are identical. We do so by comparing a proper 2-time average with the case when the time-homogeneous assumption is forcefully made. We show these differences by analyzing three examples: a DLE with a single τ-delay process and a GLE with an exponential memory and one with a power law memory, respectively, of the form

$$\phi(t) = b\mathrm{e}^{-bt}, \qquad \phi(t) = \frac{\alpha(\alpha t)^{\nu-1}}{\Gamma(\nu)}. \tag{32}$$

The Fourier characteristics and the time dependence of one-time averages of these three non-Markoffian Langevin equations have been recently studied by the present authors in [43]. The respective Green's functions are

$$\begin{aligned}
\lambda(t) &= \sum_{k=0}^{\infty} \frac{(-\gamma)^k}{k!}(t - k\tau)^k \Theta(t - k\tau),\\
\lambda(t) &= \mathrm{e}^{-bt/2}[\cos(\Omega t) + (b/2\Omega)\sin(\Omega t)],\\
\lambda(t) &= E_{\nu+1}(-\gamma\alpha^{\nu}t^{\nu+1}),
\end{aligned} \tag{33}$$

with $\Omega = \sqrt{\gamma b - b^2/4}$ and $E_{\upsilon}(z) = \sum_{n=0}^{\infty} z^n/\Gamma[n\upsilon + 1]$ the one-parameter Mittag-Leffler function of index υ [62, 63]. While the Green's function for the DLE process and for the GLE with power law memory have also a parameter region with unstable dynamics where $\lambda(t)$ grows without bounds, all three cases possess a monotonic and a (damped) oscillatory regime [43].

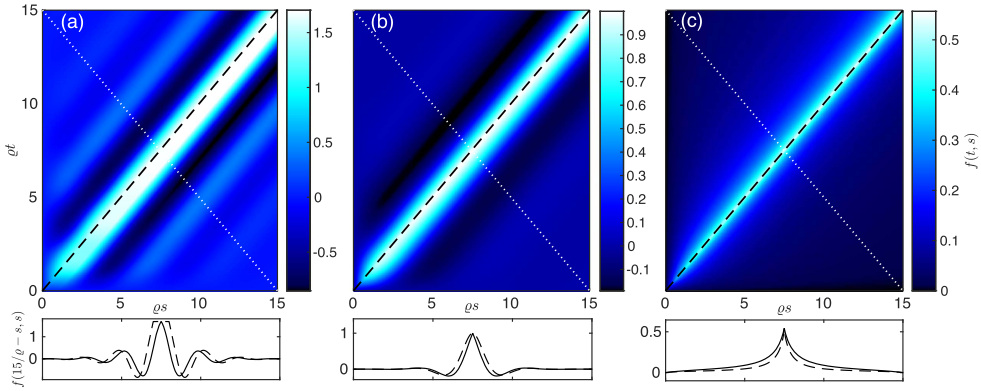

Figure 4. Correct and incorrect covariance, respectively, $f(t, s)$ in equation (34) and $f^*(t, s)$ in equation (35) of the processes governed by the single delay DLE (a), and the GLE with exponential memory (b) or algebraic memory (c). The incorrect covariance is obtained by asuming that the process is time-homogeneous. Time is normalized with respect to the coherence time scale $1/\rho$, which is, respectively, $\varrho = \sqrt{\gamma/\tau}$ for the single delay, $\varrho = \sqrt{\gamma b}$ for the exponential memory, and $\varrho = \alpha(\gamma/\alpha)^{1/(\nu+1)}$ for the algebraic memory. The parameters in the three cases are as follows: $\gamma\tau = 1$ in panel (a), $b/\gamma = 1$ in panel (b), and $\nu = -0.2$ with $\gamma/\alpha = 1$ in panel (c). Each panel is divided into two regions by the line $t = s$. For times $t > s$ the covariance is calculated using the expression (34) while for times $s < t$ the improper time-homogeneous expression (35) is used. The values of $f(t, s)$ and $f^*(t, s)$ along the dotted line orthogonal to the $t = s$ line are plotted underneath each main panel, that is $f(15/\rho - s, s)$ and $f^*(15/\rho - s, s)$ for $0 \leqslant s \leqslant 15/\rho$. In the bottom panels the solid-line represents the covariance from equation (34), while the dashed-line represents the covariance from equation (35).

We analyze the three examples by computing the covariance $\langle x(t)x(s)\rangle - \langle x(t)\rangle\langle x(s)\rangle$ where the mean is simply $\langle x(t)\rangle = \int_{-\infty}^{+\infty} \mathrm{d}x\, x\, W_1(x, t)$, and $\langle x(t)x(s)\rangle = \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y\, xy\, W_2(x, t; y, s) = \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{+\infty} \mathrm{d}y\, xy\, Q_{1|1}(x, t|y, s) W_1(y, s)$. Explicit use of equations (24) and (27) allows us to write the dimensionless covariance for $t > s$

$$f(t, s) = \frac{\langle x(t)x(s)\rangle - \langle x(t)\rangle\langle x(s)\rangle}{\sigma^2} = \int_0^s \lambda(t - r)\lambda(s - r)\mathrm{d}r, \tag{34}$$

and a corresponding expression with s interchanged with t if $t < s$. Equation (34) is valid both for a DLE with zero history and a GLE, and reduces to the 1-time $\langle x^2(t)\rangle = \nu(t)$ when $t = s$. If we were to assume that the process is homogeneous, i.e. $Q_{1|1}(x, t|y, s) = P_1(x, t - s|y, 0)$, the computation of $\langle x(t)x(s)\rangle$ for $t > s$ would only require the 1-time probability distribution and would lead to

$$f^*(t, s) = \lambda(t - s) \int_0^s \lambda^2(r)\mathrm{d}r, \tag{35}$$

with the corresponding expression with s interchanged with t if $t < s$. In figure 4 we plot the covariance from equations (34) and (35) for our three non-Markov processes in the top panels, with the former displayed only for values $t > s$, while the latter only for values $t < s$ (each respective covariance is specular around the diagonal). When $t = s$ the expressions coincide since $Q_{1|1}(x, t|y, s) = \delta(x - y)$ and the covariance reduces to a 1-time average.

When the Green's function decays to zero at long times for $t \gg s$ or $s \gg t$ the two expressions also coincide. Moving away from the diagonal is where the greater discrepancies between the correct covariance in (34) and the incorrect time-homogeneous one in (35) appear. We investigate those differences for the delay case and for the exponential memory when $\lambda(t)$ is in the oscillatory regime, whereas we study the case of the algebraic memory in the monotonic regime.

We first analyze the delay case by using the alternate expression for one-delay Green's function in equation (33) which is given by [64]

$$\lambda(t) = \sum_k \frac{\mathrm{e}^{\mathcal{Q}_k t}}{1 + \mathcal{Q}_k \tau}, \tag{36}$$

where $\mathcal{Q}_k = \mathcal{L}_k(-\gamma\tau)/\tau$ where $\mathcal{L}$ is the (multivalued) Lambert function that satisfies the relation $\mathcal{L}(z)\exp[\mathcal{L}(z)] = z$. In the stable regime for which $\Re(\mathcal{Q}_k) < 0$ for all k, if we look at values of $f(t, s)$ along a line orthogonal to the diagonal, say $t = \chi - s$, we can determine the expected rate of decay by taking only the complex conjugate pair of $\mathcal{Q}_k$ with the smallest negative real part, which we indicate for convenience with the index $k = 0$. We have in fact that for $t \geqslant s$, which implies $0 \leqslant \rho s \leqslant \chi/2$, $f(t, s) \simeq \{\exp[2\mathcal{Q}_0 \ (\chi\rho/2 - \rho s)/\sqrt{\gamma\tau}] - \exp(\mathcal{Q}_0\chi\rho/\sqrt{\gamma\tau})\}/[(1 + \mathcal{Q}_0\tau)^2 \ (-2\mathcal{Q}_0)] + \text{c.c.}$ With the parameter choice $\gamma\tau = 1$ giving the value $|Re(\mathcal{Q}_0)| = 0.32/\tau$, the approximate decay of $f(t, s)$ when $\chi = 15/\rho$ in the bottom panel of figure 4(a) matches the observed damping of the oscillations. A similar analysis shows that also equation (35) decays exponentially with the same rate.

The above analysis is useful to determine approximately the exponential decay, but is unsuccessfull to describe the great disparity as s approaches χ, and in particular it does not provide explanations as to why a nearly flat shape appears in equation (35). To explain this feature one relies upon using the top expression in (33) for $\lambda(t)$ for which $\lambda(t) = 1$ for $t \leqslant \tau$. The choice of $\rho s = \chi/2$, with χ sufficiently larger than 0, makes $\int_0^s \mathrm{d}r\ \lambda^2(r)$ nearly independent of s. In other words when the variance of the process $\nu(t)$ has already reached its saturating value, and $\chi - s \leqslant \tau$ the s-dependence of equation (35) is suppressed explaining the difference between the rapid drop of $f(t, s)$ versus the flat-top shape of $f^*(t, s)$.

The analysis with a complex exponential for $\lambda(t)$ can also be done for the case of the exponential memory with the difference that there is only decaying exponential in this case. The decay rate of $2/b$ in equation (33) corresponds to $f(t, s) \simeq \exp[-\sqrt{b/\gamma}\,(\rho\chi/2 - \rho s)]$ along the line $t = 15/\rho - s$, matching the damping of the oscillations in the bottom panel of figure 4(b). This explains the apparent difference in decay between the delay and the exponential memory case where $f(t, s)$ reduces to zero on a time scale approximately twice as fast for panel (b) versus panel (a).

For the algebraic memory we have chosen to display the case where differences between (34) and (35) are most evident, which occurs mainly in the monotonic regime. In contrast to the delay and exponential memory case the bottom panel of figure 4(c) shows the interesting feature that $f(t, s)$ decays faster than $f^*(t, s)$. It is also evident that no decay characteristic time-scale emerges at the bottom of panel (c) given the power law nature of the GLE process with algebraic memory.

Although we have focused only on the 2-time covariance, comparisons for higher multi-time averages can also be made directly from $\langle x(t)x(s)\rangle$. For Gaussian distributions it is known in fact that [49] $\langle[x(t^{(1)}) - \langle x(t^{(1)})\rangle] \ldots [x(t^{(2\ell)}) - \langle x(t^{(2\ell)})\rangle]\rangle = \sum\prod\langle[x(t^{(i)}) - \langle x(t^{(i)})\rangle][x(t^{(j)}) - \langle x(t^{(j)})\rangle]\rangle$ for any even central moment (note that ℓ is a positive integer), where the notation $\sum\prod$ indicates a sum over all distinct pairs, ij, of the 2ℓ terms On the other hand for any odd central moment, that is when computing $2\ell - 1$ terms, we have

$\langle [x(t^{(1)}) - \langle x(t^{(1)})\rangle] \ldots [x(t^{(2\ell-1)}) - \langle x(t^{(2\ell-1)})\rangle]\rangle = 0$. For example the three time average is explicitly obtained in terms of the lower order moments as

$$\begin{aligned}\langle x(t)x(s)x(r)\rangle = {} & \langle x(t)\rangle[\langle x(s)x(r)\rangle - \langle x(s)\rangle\langle x(r)\rangle] \\ & + \langle x(s)\rangle[\langle x(t)x(r)\rangle - \langle x(t)\rangle\langle x(r)\rangle] \\ & + \langle x(r)\rangle[\langle x(t)x(s)\rangle - \langle x(t)\rangle\langle x(s)\rangle] + \langle x(t)\rangle\langle x(s)\rangle\langle x(r)\rangle, \end{aligned} \tag{37}$$

which clearly shows that to calculate $\langle x(t)x(s)x(r)\rangle$ is sufficient to know $\langle x(t)\rangle$ and $\langle x(t)x(t')\rangle$. The formalism presented here can thus give any even order moments higher than 2.

6. Generalization of the Green–Kubo formula

We now make use of the formalism developed in the previous sections and present a generalization of the Green–Kubo formula relating the spatial diffusion coefficient to the velocity autocorrelation function. A recent study [65] has elucidated the relation between the long-time temporal dependence of the velocity autocorrelation function and the anomalous scaling behavior of the MSD in systems possessing scale invariant dynamics. In this section we are interested in presenting exact time-dependent expressions of the MSD in terms of the Green's function $\lambda(t)$ when the velocity associated to a random variable is governed by a DLE or GLE, focusing in particular on systems that do not obey the fluctuation–dissipation relation.

To compute the general expression for the MSD we evaluate the velocity autocorrelation function $\langle v(t_2)v(t_1)\rangle = \int_{-\infty}^{+\infty} \mathrm{d}v_2 \int_{-\infty}^{+\infty} \mathrm{d}v_1\, v_2 v_1\, Q_{1|1}(v_2, t_2|v_1, t_1) W_1(v_1, t_1)$. The MSD $\mathrm{d}^2(t) = \langle [x(t) - \langle x(t)\rangle]^2\rangle = \left\langle \left[\int_0^t \mathrm{d}s\, v(s) - \left\langle \int_0^t \mathrm{d}s\, v(s)\right\rangle\right]^2 \right\rangle$ can then be expressed in terms of the Green's function as

$$\mathrm{d}^2(t) = \sigma^2 \int_0^t \mathrm{d}t_2 \int_0^{t_2} \mathrm{d}t_1 \int_0^{t_1} \mathrm{d}s\, \lambda(t_2 - s)\lambda(t_1 - s), \tag{38}$$

where now σ^2 has units of velocity squared over time. In equation (38) we have used the fact that $\lambda(z) = 0$ when $z < 0$ to change the limit of integration of the second integral from t to t_2 on the right-hand side. Notice that the traditional simplification of taking $W_1(v_1, t_1)$ to be the stationary distribution $W_1(v_1, t_1 \to +\infty)$ [65] has not been employed here.

In the delay case it is possible to integrate equation (36) explicitly and show that at long times $\mathrm{d}^2(t) \simeq \sigma^2 t \sum_{k,l} [\mathcal{Q}_k(\mathcal{Q}_k + \mathcal{Q}_l)(1 + \mathcal{Q}_k\tau)(1 + \mathcal{Q}_l\tau)]^{-1}$. To determine the diffusion coefficient we see that $\sum_{k,l}\{[\mathcal{Q}_k(1 + \mathcal{Q}_k\tau)(1 + \mathcal{Q}_l\tau)]^{-1} + [\mathcal{Q}_l(1 + \mathcal{Q}_k\tau)(1 + \mathcal{Q}_l\tau)]^{-1}\} = \{\sum_n[\mathcal{Q}_n(1 + \mathcal{Q}_n\tau)]^{-1}\}^2$, and that the Laplace transform of equation (36) can be written alternatively as [42] $\tilde{\lambda}(\epsilon) = [\epsilon + \gamma\exp(-\epsilon\tau)]^{-1}$ or $\tilde{\lambda}(\epsilon) = \sum_k [(\mathcal{Q}_k - \epsilon)(1 + \mathcal{Q}_k\tau)]^{-1}$ where ϵ is the Laplace variable. Since $\tilde{\lambda}(0) = \gamma^{-1}$ it is then straightforward to write the long-time dependence of the MSD as $\mathrm{d}^2(t) \simeq \sigma^2 t/\gamma^2$ by noticing that $\tilde{\lambda}(0) = \gamma^{-1}$. This temporal dependence is visible early on in the overdamped regime ($\gamma\tau < e^{-1}$), but appears at increasingly longer time scales as the system approaches the unstable regime ($\gamma\tau \to \pi/2$) as depicted in figure 5(a). A simple calculation shows that also in the exponential memory case the long time MSD converges to $\sigma^2 t/\gamma^2$.

At intermediate times the MSD dynamics for the delayed system follows the intuitive understanding that the system becomes diffusive depending on the decay time of the Green's function, namely the faster $\lambda(t)$ decays the earlier the diffusive regime sets in. This is visible in figure 5(a) where we have plotted the MSD as $\gamma\tau$ decreases. As the magnitude of $\mathcal{Q}_0$ increases with a decrease in $\gamma\tau$ one notices that the system goes from the overdamped to the

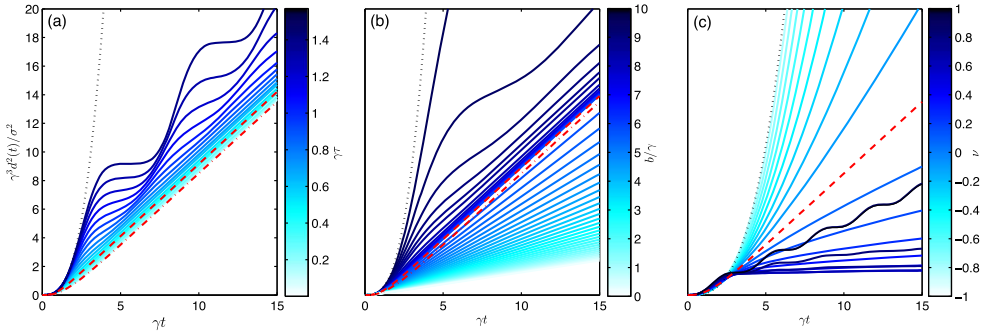

Figure 5. MSD dynamics in equation (38) when the Langevin equation in (3) or (30) governs the change over time of the velocity. Panels (a)–(c) represent, respectively, the delayed Langevin case with one delay (a), the GLE with exponential memory, and (c) the GLE with power law memory; the corresponding Green's function expressions are presented in equation (33). In all three panels the MSD in the limit $\gamma = 0$, i.e. $\sigma^2 t^3/3$, is drawn with a dotted line representing the Wiener dynamics, which is common to all curves at short times. In all three panels the dashed red line separates the top curves where the Green's function is underdamped from the bottom curves where the Green's function is overdamped. In panels (a)–(c) that corresponds, respectively, to the value $\gamma\tau = \mathrm{e}^{-1}$, $\gamma/b = 1/4$ and $\nu = 0$. In panel (a) the dash-dotted line is the $\tau = 0$ case, which represents the MSD for the Ornstein–Uhlenbeck case whose analytic expression can be written as $\sigma^2[2\gamma t + 4\exp(-\gamma t) - \exp(-2\gamma t) - 3]/(2\gamma^3)$. The same exact expression is drawn as a dashed line in panel (c) ($\nu = 0$ case) as it separates the underdamped from the overdamped regime.

underdamped regime as $\gamma\tau \to 0$ (for any non-zero γ) with the diffusive linear dependence being approached at increasingly earlier times. The lowest lying dashed–dotted curve has $\tau = 0$, which corresponds to an Ornstein–Uhlenbeck process with rate γ.

The dynamics of the MSD for the exponential memory follows a qualitative similar trend in the underdamped regime. There oscillations gets progressively reduced as the decay rate of the Green's function, in this case b, increases. With the choice of γ/b as our dimensionless parameter we see in figure 5(b) that with smaller values of γ/b the diffusive regime appears at earlier times. The trend in the overdamped regime, i.e. below the dashed line, is however different from the delay case. Although the MSD eventually reaches the common diffusive limit, $\sigma^2 t/\gamma^2$, for any non-zero γ, a reduction in γ/b pushes the curves further and further from the dashed line. Numerically we have verified that the location where the curvature is maximal, i.e. the γt value where the third derivative of the MSD equals zero, grows unbounded as γ/b approaches 0.

The MSD for the power law memory in figure 5(c) has, on the other hand, a very different dynamics from the MSD in panels (a) and (b). The reason is the power law nature of the Green's function. In contrast to both the delay and the exponential memory, the qualitative dynamics shown in figure 5(c) depend entirely on the algebraic parameter ν; the scaling coefficient $\gamma\alpha^\nu$, from equation (33), affects only the scaling of time [43]. In the undamped regime of the algebraic memory, for which ν is between 0 and 1, the qualitative behavior of the MSD at long times is similar to that of both the delay and exponential memories as it displays diffusive dynamics with the MSD proportional to t. The proportionality constant is dependent on the value of ν. An increase in ν from 0 leads to an initial decrease in the value of the diffusion coefficient, however, as ν is increased further the coefficient eventually increases. This non-monotonic dependence is evident in figure 5(c) below the dashed line.

The value of ν that minimizes the diffusion constant strongly depends on the scaling coefficient.

By contrast, the dynamics in the over-damped regime is superdiffusive. This is evident through an analytical examination of the well-known long-time approximation of the Mittag-Leffler function [63], i.e., $\lambda(t)$, valid for non-integer values of ν in the range $|\nu| < 1$,

$$\lambda(t \to \infty) = -\sum_{n=1}^{p} \frac{1}{\Gamma[1 - n(\nu + 1)]} \left[\frac{-1}{\gamma \alpha^{\nu} t^{\nu+1}} \right]^{n} + O(t^{-p(\nu+1)}). \tag{39}$$

Here p is an integer and $O(...)$ is the standard order function which indicates the dominant term in the asymptotic limit. The leading term of equation (39) is proportional to $1/t^{\nu+1}$. In the over-damped regime, being of an order greater than -1, it dominates the integrals of equation (38) when $t \to \infty$. Upon the evaluation of equation (38), with an exclusive focus on the long-time behavior, the MSD is given approximately by $\mathrm{d}^2(t) \propto t^{1-2\nu}$ thus it approaches the standard Weiner dynamics, i.e., t^3, as ν approaches -1. This leading order analytic result qualitatively describes the long-time dependence of the MSD on ν that is depicted in figure 5(c) above the dashed line. The extreme short-time correlations in the overdamped parameter regime of algebraic memory results in a strongly correlated forcing term and, subsequently, strongly correlated motion.

7. Conclusions

Stochastic delay differential equations have a long tradition. Since their original introduction to study the macrodynamic theory of business cycles [66, 67], they have been employed across a multitude of scientific areas. In recent years they have been used in particular to study the lag dynamics and synchronizability of networks [68–73] whereby each node may interact with any other node with different coupling strength and with different delay. Although our results are presented for the case of a multi-delayed stochastic system in one dimension, the approach can be employed to study the dynamics of multidimensional Langevin equations with multiple delays given by $\dot{\mathbf{x}}(t) = -\sum_i \mathbf{A}_i \mathbf{x}(t - \tau_i) + \sigma\eta(t)$. Our FPE may thus represent a useful limiting benchmark to unravel synchronizability issues in certain types of networks.

The specific focus of our analysis on constructing *bona fide* FPEs for systems obeying a DLE has an important general applicability. It provides the ability to analyze systems governed by time non-local harmonic forces and subject to boundary conditions. Already for the simpler case of time local harmonic forces knowing the exact functional dependence of the 1- and 2-time probability distributions in infinite space is, in most cases, not sufficient to construct the corresponding distributions in presence of absorbing, reflecting or periodic boundaries. Techniques such as the method of images cannot be employed in situations where the space is not homogenous [74, 75]. Two are the suggested solutions to circumvent this difficulty. By finding the exact dynamics of the Langevin equation that explicitly incorporate the presence of boundaries, it would be possible to construct the associated probability distribution via characteristic functionals as suggested by Budini and Caceres [50]. When such an exact Langevin solution is lacking, the other approach consists of supplementing the appropriate boundary conditions at the probability level and attempting to build an analytic or numerical solution of the FPE. Determining a *bona fide* FPE thus promise to open up opportunities to tackle a variety of boundary situations in systems governed by DLEs, which are currently accessible only via stochastic simulations.

In this study we have exploited the linearity of the DLE to generate an equivalent memoryless Langevin equation with colored Gaussian noise via the so-called time-

convolutionless transform. We have then constructed the appropriate FPE for the 1- and 2-time probability distribution for the DLE. We have contrasted our FPE formalism with the one proposed in the past where the dynamics of the 1-time probability distribution is coupled to the moments of the 2-time probability distribution. Although we have found an appropriate ansatz that allows to solve exactly the proposed non-closed FPE, we have shown that it can only describe one-time averages and does not provide a *bona fide* representation of the random process.

Beyond the DLE we have shown that the 1- and 2-time FPE construction procedure can be used essentially unaltered for the case of a GLE. In this respect our results extend known procedures to build 1- and 2-time FPE for a GLE in which fluctuation–dissipation is valid to cases in which it is not and noise is white. The derived analytic expressions for W_1 and W_2 probability distribution have allowed us to show that both DLEs and GLEs are non-Markov time-inhomogeneous processes. Time-inhomogeneity is characteristic of ageing systems for which it is necessary to distinguish between preparation and observation times. The time-dependent dynamics cannot in fact be evinced only from W_1, but requires the knowledge of the conditional probability distribution $Q_{1|1}$.

By taking a specific two-delay system we have presented in detail how the dynamics of the probability distribution differ if observation starts at time $t = 0$ versus some later time. But we have also shown that, even when observation and preparation times coincide, inaccuracies still emerge in evaluating multi-time averages if one were to assume that the process is time-homogeneous. To show this aspect explicitly we have also analyzed and compared the effects of assuming time-homogeneity when computing the covariance for a one-delay DLE and two GLEs, respectively, with an exponential memory and a power law memory.

Finally, application to the case in which the Langevin equation describes the stochastic dynamics of the velocity variable rather than position has allowed us to generalize the Green–Kubo formalism to situations in which a system does not obey the fluctuation dissipation relation.

Acknowledgments

This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) UK Grant No. EP/I013717/1. LG acknowledges support from the EPSRC grant EP/I013717/1. All underlying data are provided in full within this paper.

Appendix A. Non-closed FPE for delay processes and its solution

One may construct a FP equation for the 1-time probability distribution associated with (3) using a simple generalization of the procedure presented in [36, 52] for the case of one delay. With the substitution of equation (3) into equation (12) one has

$$\frac{\partial}{\partial t}W_1(x,\,t) = -\frac{\partial}{\partial x}\left\langle\left[-\sum_{i=1}^{N}\gamma_i x(t-\tau_i) + \sigma\eta(t)\right]\delta(x - x(t))\right\rangle. \tag{A.1}$$

Defining the fixed variable, x_{τ_i} corresponding to the stochastic variable, $x(t-\tau_i)$ and using the following two properties, $\langle f(x(t))\delta(x - x(t))\rangle = \langle\int \mathrm{d}x' f(x(t))\delta(x - x(t))\delta(x' - x(t'))\rangle$ and $\int \mathrm{d}z\, z \quad \delta(z - s) = \int \mathrm{d}s\, s\, \delta(z - s)$, one notices that $\langle y_{\tau_i}\delta(x - y)\rangle = \langle\int \mathrm{d}x_{\tau_i} x_{\tau_i}\, \delta(x - y)\delta(x_{\tau_i} - y_{\tau_i})\rangle = \int \mathrm{d}x_{\tau_i}\, x_{\tau_i} W_2(x,\,t;\,x_{\tau_i},\,t-\tau_i)$. After rewriting equation (A.1) as

$$\frac{\partial}{\partial t}W_1(x,\,t) = \frac{\partial}{\partial x}\sum_{i=1}^{N}\gamma_i\int_{-\infty}^{\infty}\mathrm{d}x_{\tau_i}\,x_{\tau_i}\,W_2(x,\,t;\,x_{\tau_i},\,t-\tau_i) - \frac{\partial}{\partial x}\langle\sigma\eta(t)\delta(x-x(t))\rangle, \tag{A.2}$$

one is left to evaluate the average on the right-hand side using Novikov's theorem [53], which states

$$\begin{aligned}\sigma\langle\{\eta(t)\delta(x-x(t))\}\rangle &= \sigma\int_0^t \mathrm{d}s\langle\eta(t)\eta(s)\rangle\left\langle\frac{\delta[\delta(x(t)-x)]}{\delta\eta(s)}\right\rangle\\ &= -\sigma\int_0^t \mathrm{d}s\,\delta(t-s)\left\langle\frac{\partial}{\partial x}\left(\left[\frac{\delta x(t)}{\delta\eta(s)}\right]_{x(t)=x}\delta(x(t)-x)\right)\right\rangle\\ &= -\sigma\frac{\partial}{\partial x}\left\langle\left[\frac{\delta x(t)}{\delta\eta(t)}\right]_{x(t)=x}\delta(x(t)-x)\right\rangle.\end{aligned} \tag{A.3}$$

Here we have used the delta correlated property of the noise and the functional chain rule to obtain the average in terms of the functional derivative of x with respect to the noise term. For the linear Langevin equation, equation (3), one can calculate these functional derivatives exactly [27, 28],

$$\frac{\delta x(t)}{\delta\eta(t)} = \frac{\sigma}{2}. \tag{A.4}$$

Substituting this expression into (A.3) we obtain

$$\sigma\langle\{\eta(t)\delta(x-x(t))\}\rangle = -\frac{\sigma^2}{2}\frac{\partial}{\partial x}W_1(x,\,t) \tag{A.5}$$

which in turn can be substituted in equation (A.2) to give the desired FP

$$\frac{\partial}{\partial t}W_1(x,\,t) = \frac{\partial}{\partial x}\sum_{i=1}^{N}\gamma_i\int_{-\infty}^{\infty}\mathrm{d}x_{\tau_i}\,x_{\tau_i}\,W_2(x,\,t;\,x_{\tau_i},\,t-\tau_i) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}W_1(x,\,t), \tag{A.6}$$

where x_{τ_i} represents $x(t-\tau_i)$.

Equation (A.6) was already derived in [38], but it was never solved exactly. Writing FPEs for higher order n-time probability distributions does not allow one to close the system of equations since it only creates an infinite hierarchy of equations with W_n depending on W_{n+1} [76]. There is however a way to find the exact solution once the initial distribution, $W_1(x,\,0) = W_0(x)$, and the joint distribution $W_2(x,\,t;\,x',\,t')$ for $t > t'$ and $t' \geqslant 0$ are specified, while for $t' \in [-\tau_N,\,0]$ the dynamics is deterministic and dictated by the history function $\phi(t)$ such that $W_2(x,\,0;\,x',\,t') = W_0(x)W_1(x',\,t')$ and $W_1(x',\,t') = \delta(x'-\phi(t'))$.

For an equivalent representation of the Langevin equation (3), the initial distribution has to take the form of a Dirac-delta function, $W_0(x) = \delta(x-x_0)$, and the joint distribution needs to assign the deterministic dynamics given by the history function $\phi(t)$ for $t' < 0$, i.e. $W_2(x,\,0;\,x',\,t') = W_0(x)W_1(x',\,t')$ and $W_1(x',\,t') = \delta(x'-\phi(t'))$ for $t' \in [-\tau_N,\,0]$.

To find the solution $W_1(x,\,t)$, given the underdetermined nature of equation (A.6), one proposes an ansatz in terms of the moments of the distribution and supplement information about those moments from the Langevin equation. Such a procedure has been exploited to find an exact solution at steady state for the case of a single delay dynamics [39, 77]. We now show that a proper ansatz allows one to extend the procedure to the case of multiple delays and for the full time-dependence.

We begin by substituting the marginal relationship, $W_1(x,\,t) = \int W_2(x,\,t;\,x_{\tau_i},\,t-\tau_i)\mathrm{d}x_{\tau_i}$ into equation (A.6) and obtain

$$\int_{-\infty}^{\infty} \mathrm{d}x_{\tau_1} \int_{-\infty}^{\infty} \mathrm{d}x_{\tau_2} \ \ldots \int_{-\infty}^{\infty} \mathrm{d}x_{\tau_N} \left[\frac{\partial}{\partial t} W_N(x, t; x_{\tau_1}, t-\tau_1; \ldots, x_{\tau_N}, t-\tau_N) \right.$$
$$- \frac{\partial}{\partial x} \sum_{i=1}^{N} \gamma_i x_{\tau_i} W_N(x, t; x_{\tau_1}, t-\tau_1; \ldots, x_{\tau_N}, t-\tau_N)$$
$$\left. - \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} W_N(x, t; x_{\tau_1}, t-\tau_1; \ldots, x_{\tau_N}, t-\tau_N) \right] = 0. \tag{A.7}$$

Naming, k and k_i, as the associated Fourier components of x and x_{τ_i}, we can write the Fourier transform of this probability distribution as

$$\hat{W}_N(k, t; k_1, t-\tau_1; \ldots, k_N, t-\tau_N) = \int_{-\infty}^{\infty} \mathrm{d}x \int_{-\infty}^{\infty} \mathrm{d}x_{\tau_1} \ldots$$
$$\times \int_{-\infty}^{\infty} \mathrm{d}x_{\tau_N} \, \mathrm{e}^{-\mathrm{i}kx} \prod_{i=1}^{N} \mathrm{e}^{-\mathrm{i}k_i x_{\tau_i}} W_N(x, t; x_{\tau_1}, t-\tau_1; \ldots, x_{\tau_N}, t-\tau_N). \tag{A.8}$$

Multiplying equation (A.7) by $\mathrm{e}^{-\mathrm{i}kx}$ and integrating over x we arrive at the transformed equation

$$\left[\frac{\partial}{\partial t} \hat{W}_N(k, t; k_1, t-\tau_1; \ldots, k_N, t-\tau_N) + k \sum_{i=1}^{N} \gamma_i \frac{\partial}{\partial k_i} \hat{W}_N(k, t; k_1, t-\tau_1; \ldots, k_N, t-\tau_N) \right.$$
$$\left. - \frac{\sigma^2}{2} k^2 \hat{W}_N(k, t; k_1, t-\tau_1; \ldots, k_N, t-\tau_N) \right]_{\forall k_i = 0} = 0, \tag{A.9}$$

where the specific constraint of $k_i = 0$ for each i means that $\hat{W}_1(k, t)$ is simply the Fourier transform of the marginal distribution $W_1(x, t)$. The transformed equation (A.9) has the same form as a multivariate Ornstein–Uhlenbeck process for which we can propose the general ansatz in Fourier space [78]

$$\hat{W}_N(k, t; k_1, t-\tau_1; \ldots, k_N, t-\tau_N) = \exp\left\{ -\mathrm{i}k\mu(t) - \mathrm{i} \sum_{i=1}^{N} k_i \mu(t-\tau_i) \right.$$
$$\left. - \frac{\sigma^2}{2} \left[k^2 c(t, t) + k \sum_{i=1}^{N} k_i c(t, t-\tau_i) - \sum_{i=1}^{N} \sum_{j=1}^{N} k_i k_j c(t-\tau_i, t-\tau_j) \right] \right\}. \tag{A.10}$$

Equation (A.10) is a multivariate normal distribution with mean $\mu(t)$ and covariance $c(t, t')$. Substituting (A.10) into (A.9), setting $k_i = 0$ for all i, gives the following two conditions

$$\dot{\mu}(t) = -\sum_{i=1}^{N} \gamma_i \mu(t-\tau_i), \qquad \dot{c}(t, t) = -2 \sum_{i=1}^{N} \gamma_i c(t, t-\tau_i) + \sigma^2. \tag{A.11}$$

To obtain an explicit expression for $\hat{W}_1(k, t)$ one needs to solve the two differential equation (A.11) relating the evolution of the mean, variance and covariance of the system. In order to proceed we have to return to the Langevin equation equation (3), and examine the evolution of its mean and variance

$$\frac{\partial}{\partial t}\langle x(t)\rangle = -\sum_{i=1}^{N}\gamma_i\langle x(t-\tau_i)\rangle,$$
$$\frac{\partial}{\partial t}[\langle x^2(t)\rangle - \langle x(t)\rangle^2] = -2\sum_{i=1}^{N}\gamma_i[\langle x(t)x(t-\tau_i)\rangle - \langle x(t)\rangle\langle x(t-\tau_i)\rangle] + \sigma^2. \tag{A.12}$$

It is thus evident that the mean and variance (A.12) of the Langevin equation (3) obey the same dynamics of the mean and variance of the ansatz distribution (A.11). Therefore we can use equation (5) and the properties of the Gaussian noise to find the following explicit expressions

$$\langle x(t)\rangle = x_0\lambda(t) + \Psi_d(t)\langle x^2(t)\rangle - \langle x(t)\rangle^2 = \int_0^t \mathrm{d}s\ \lambda^2(s).$$

Note that the expression for the mean is dependent on the initial condition of the system, x_0. Therefore substituting these averages into equation (A.10), and settting $k_\tau = 0$, will give a Gaussian that satisfies the FPE and represents the initially conditioned distribution, $P_1(x, t|x_0, 0)$, in the Fourier domain

$$\hat{P}_1(k, t|x_0, 0) = \exp\left\{-\mathrm{i}k\left[x_0\lambda(t) - \sum_{i=1}^{N}\gamma_i\int_{-\tau_i}^{0}\mathrm{d}s\ \lambda(t-\tau_i-s)\phi(s)\right] - \frac{1}{2}k^2\int_0^t \mathrm{d}s\ \lambda^2(s)\right\} \tag{A.13}$$

From the general relation between W_n and P_n it then follows that

$$W_1(x, t) = \left[2\pi\sigma^2\int_0^t \mathrm{d}s\ \lambda^2(s)\right]^{-1/2}\int_{-\infty}^{\infty}\mathrm{d}z\ W_0(z) \times \exp\left\{-\frac{\left(x - x_0\lambda(t) + \sum_{i=1}^{N}\gamma_i\int_{-\tau_i}^{0}\mathrm{d}s\ \lambda(t-\tau_i-s)\phi(s)\right)^2}{2\sigma^2\int_0^t \mathrm{d}s\ \lambda^2(s)}\right\}. \tag{A.14}$$

In the respective limits $\tau_i = 0$ (and $\phi(t) = 0$) or $\gamma_i = 0$ for each i, the delay Langevin equation (3) reduces to the one describing the Ornstein–Uhlenbeck process or the Wiener process. In these cases $\lambda(t) = \mathrm{e}^{-\sum_{i=1}^{N}\gamma_i t}$ and $\lambda(t) = 1$, respectively. The expression (A.14) is identical to the one obtained via characteristic functionals [42].

Appendix B. Derivation of the 1- and 2-time FPEs

We derive the closed FPEs for the 1- and 2-time probability distributions using the same method as in appendix A but with the time-convolutionless Langevin equation (7).

We begin by taking the derivative with respect to time of $W_1(x, t) = \langle\delta(x - x(t))\rangle$ and $W_2(x, t; x', t') = \langle\delta(x - x(t))\delta(x' - x(t'))\rangle$, and substituting in equation (7), to give

$$\frac{\partial}{\partial t}W_1(x, t) = -\frac{\partial}{\partial x}\langle[-A(t)x(t) + B_d(t) + \sigma\xi(t)]\delta(x - x(t))\rangle,$$
$$\frac{\partial}{\partial t}W_2(x, t; x', t') = -\frac{\partial}{\partial x}\langle[-A(t)x(t) + B_d(t) + \sigma\xi(t)]\delta(x - x(t))\delta(x' - x(t'))\rangle. \tag{B.1}$$

Using the property $\langle y\delta(x-y)\rangle = x\langle\delta(x-y)\rangle$ we can immediately evaluate the first terms on the right-hand side

$$\begin{aligned}\frac{\partial}{\partial t}W_1(x,t) &= -\frac{\partial}{\partial x}[-A(t)x + B_d(t)]W_1(x,t) - \frac{\partial}{\partial x}\langle\sigma\xi(t)\delta(x-x(t))\rangle,\\ \frac{\partial}{\partial t}W_2(x,t;x',t') &= -\frac{\partial}{\partial x}[-A(t)x + B_d(t)]W_2(x,t;x',t')\\ &\quad -\frac{\partial}{\partial x}\langle\sigma\xi(t)\delta(x-x(t))\delta(x'-x(t'))\rangle.\end{aligned} \tag{B.2}$$

The remaining averages can once again be calculated using Novikov's theorem [53]

$$\begin{aligned}\sigma\langle\xi(t)\delta(x-x(t))\rangle &= \sigma\int_0^t \mathrm{d}s\langle\xi(t)\xi(s)\rangle\left\langle\frac{\delta[\delta(x(t)-x)]}{\delta\xi(s)}\right\rangle\\ &= -\sigma\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\left\langle\frac{\partial}{\partial x}\left(\left[\frac{\delta x(t)}{\delta\eta(s)}\right]_{x(t)=x}\delta(x(t)-x)\right)\right\rangle\\ &= -\sigma\frac{\partial}{\partial x}\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\left[\frac{\delta x(t)}{\delta\eta(s)}\right]_{x(t)=x}\langle\delta(x(t)-x)\rangle.\end{aligned} \tag{B.3}$$

$$\begin{aligned}&\sigma\langle\xi(t)\delta(x-x(t))\delta(x'-x(t'))\rangle\\ &\quad= \sigma\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\left\langle\frac{\delta[\delta(x(t)-x)]}{\delta\xi(s)}\delta(x(t')-x')\right\rangle\\ &\qquad+\sigma\int_0^{t'} \mathrm{d}s'\,\langle\xi(t)\xi(s')\rangle\left\langle\delta(x(t)-x)\frac{\delta[\delta(x(t')-x')]}{\delta\xi(s')}\right\rangle\\ &\quad= -\sigma\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\frac{\partial}{\partial x}\left\langle\left[\frac{\delta x(t)}{\delta\xi(s)}\right]_{x(t)=x}\delta(x(t)-x)\delta(x(t')-x')\right\rangle\\ &\qquad-\sigma\int_0^{t'} \mathrm{d}s'\,\langle\xi(t)\xi(s')\rangle\frac{\partial}{\partial x'}\left\langle\left[\frac{\delta x(t')}{\delta\xi(s')}\right]_{x(t')=x'}\delta(x(t)-x)\delta(x(t')-x')\right\rangle\\ &\quad= -\sigma\frac{\partial}{\partial x}\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\left[\frac{\delta x(t)}{\delta\xi(s)}\right]_{x(t)=x}\langle\delta(x(t)-x)\delta(x(t')-x')\rangle\\ &\qquad-\sigma\frac{\partial}{\partial x}\int_0^{t'} \mathrm{d}s'\,\langle\xi(t)\xi(s')\rangle\left[\frac{\delta x(t')}{\delta\xi(s')}\right]_{x(t')=x'}\langle\delta(x(t)-x)\delta(x(t')-x')\rangle.\end{aligned} \tag{B.4}$$

The functional derivative can be calculated exactly as

$$\frac{\delta x(t)}{\delta\xi(s)} = \exp\left\{\int_s^t \mathrm{d}r\,\frac{\frac{\partial}{\partial r}\lambda(r)}{\lambda(r)}\right\} = \sigma\frac{\lambda(t)}{\lambda(s)}, \tag{B.5}$$

and substituted in equation (B.4) to give

$$\begin{aligned}\frac{\partial}{\partial t}W_1(x,\,t) &= -\frac{\partial}{\partial x}[-A(t)x + B_d(t)]W_1(x,\,t)\\ &\quad + \sigma^2\frac{\partial^2}{\partial x^2}\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\frac{\lambda(t)}{\lambda(s)}W_1(x,\,t),\\ \frac{\partial}{\partial t}W_2(x,\,t;\,x',\,t') &= -\frac{\partial}{\partial x}[-A(t)x + B_d(t)]W_2(x,\,t;\,x',\,t')\\ &\quad + \sigma^2\frac{\partial^2}{\partial x^2}\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\frac{\lambda(t)}{\lambda(s)}W_2(x,\,t;\,x',\,t')\\ &\quad + \sigma^2\frac{\partial^2}{\partial x\partial x'}\int_0^{t'} \mathrm{d}s'\,\langle\xi(t)\xi(s')\rangle\frac{\lambda(t')}{\lambda(s')}W_2(x,\,t;\,x',\,t').\end{aligned} \tag{B.6}$$

With the help of equation (11) all that remains is to evaluate

$$\int_0^t \mathrm{d}s\,\langle\xi(t)\xi(s)\rangle\frac{\lambda(t)}{\lambda(s)} = \frac{\lambda^2(t)}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{d}s\,\frac{\lambda^2(t-s)}{\lambda^2(t)}, \tag{B.7}$$

$$\int_0^{t'} \mathrm{d}s'\,\langle\xi(t)\xi(s')\rangle\frac{\lambda(t')}{\lambda(s')} = \lambda(t)\frac{\mathrm{d}}{\mathrm{d}t}\int_0^{t'} \mathrm{d}s\frac{\lambda(t-s)\lambda(t'-s)}{\lambda(t)}. \tag{B.8}$$

Substituting these expressions into equation (B.6) we obtain the desired FPEs (14) and (15).

Appendix C. Solving the FPEs

To solve the FPEs for the 1- and 2-time probability distributions we use the method of transformation of variables [79]. Here we demonstrate the method with a general FPE whose structure represents both equations (14) and (15). Consider the distribution $R(x,\,t',\,t)$ obeying

$$\frac{\partial}{\partial t}R(x,\,t',\,t) = -\frac{\partial}{\partial x}\left[\frac{\dot{a}(t)}{a(t)}x + b(t',\,t) - c(t)\frac{\partial}{\partial x}\right]R(x,\,t',\,t) \tag{C.1}$$

with $0 \leqslant t' < t$ and subject to the general initial condition, $R(x,\,t',\,t') = \delta(x-y)Z(y,\,t')$. We define the variable transformation

$$R(x,\,t',\,t) = \frac{a(t')}{a(t)}V(y(x,\,t',\,t),\,\varrho(t',\,t)), \tag{C.2}$$

where

$$y(x,\,t',\,t) = \frac{a(t')}{a(t)}x - \int_{t'}^t \mathrm{d}s\,b(t',\,s)\frac{a(t')}{a(s)}, \qquad \varrho(t',\,t) = \int_{t'}^t \mathrm{d}s\,c(s)\frac{a^2(t')}{a^2(s)}, \tag{C.3}$$

and where the choice of t' for the lower limit of the integrals is taken to satisfy the initial conditions. Substituting $R(x,\,t',\,t)$ in equation (C.1) using equation (C.2) gives the diffusion equation,

$$\frac{\partial V(y,\,\varrho)}{\partial\varrho} = \frac{\partial^2 V(y,\,\varrho)}{\partial y^2}, \tag{C.4}$$

which has the well known solution

$$V(z,\,\rho) = \frac{1}{\sqrt{4\pi\varrho}}\int_{-\infty}^{\infty} \mathrm{d}\omega\; V(\omega,\,0)\exp\left\{-\frac{(z-\omega)^2}{4\varrho}\right\}. \tag{C.5}$$

Re-expressing $V(z,\,\varrho)$ in terms of the original variables and noting that $R(x,\,t',\,t') = V(x,\,0)$ we can write the exact solution as

$$R(x,\,t',\,t) = \left[4\pi\int_{t'}^{t}\mathrm{d}s\;\frac{c(s)a^2(t)}{a^2(s)}\right]^{-1/2} Z(y,\,t') \times \exp\left\{-\frac{\left(x-\int_{t'}^{t}\mathrm{d}s\;b(t',\,s)\frac{a(t)}{a(s)}\right)^2 - y\frac{a(t)}{a(t')}\Big)^2}{4\int_{t'}^{t}\mathrm{d}s\;\frac{c(s)a^2(t)}{a^2(s)}}\right\}. \tag{C.6}$$

From this expression we can immediately obtain the solution of equation (14) for $W_1(x,\,t)$. Substituting $t' = 0$, $a(t) = \lambda(t)$, $c(t) = D(t)$ and $b(t',\,t) = B_d(t)$ into (C.6) we arrive at equation (24). To obtain the solution for $W_2(x,\,t;\,x',\,t')$ more work is required as equation (C.1) does not correspond directly with equation (15) but does correspond to the Fourier transform of (15) over x', given by

$$\frac{\partial}{\partial t}\hat{W}_2(x,\,t;\,k',\,t') = \frac{\partial}{\partial x}\left[A(t)x - B_d(t) + \mathrm{i}k'C(t',\,t) + D(t)\frac{\partial}{\partial x}\right]\hat{W}_2(x,\,t;\,k',\,t'), \tag{C.7}$$

where $\hat{W}_2(x,\,t;\,k',\,t') = \int_{-\infty}^{\infty}\mathrm{d}x'\mathrm{e}^{-\mathrm{i}k'x'}W_2(x,\,t;\,x',\,t')$. Now equations (C.7) and (14) are of the same form and thus the solution of equation (C.7) is given by equation (C.6) with $a(t) = \lambda(t)$, $c(t) = D(t)$ and $b(t',\,t) = B_d(t) - \mathrm{i}k'C(t',\,t)$. Substituting these in along with the Fourier transformed initial condition, $\hat{W}_2(x,\,t';\,k',\,t') = \mathrm{e}^{-\mathrm{i}k'x}W_1(x,\,t')$, gives

$$\hat{W}_2(x,\,t;\,k',\,t') = \int_{-\infty}^{\infty}\mathrm{d}\omega\;\mathrm{e}^{-\mathrm{i}k'\omega}\frac{W_1(\omega,\,t')}{\sqrt{4\pi\rho(t',\,t)}} \times \exp\left\{-\frac{\left[x-\omega\frac{\lambda(t)}{\lambda(t')} - \Psi_d(t) + \Psi_d(t') + \mathrm{i}k'\beta(t',\,t)\right]^2}{4\rho(t',\,t)}\right\}. \tag{C.8}$$

Here we have defined

$$\rho(t',\,t) = \int_{t'}^{t}\mathrm{d}s\;D(s)\frac{\lambda^2(t)}{\lambda^2(s)} = \frac{\nu(t)}{2} - \frac{\lambda^2(t)}{2\lambda^2(t')}\nu(t'), \tag{C.9}$$

$$\beta(t',\,t) = \int_{t'}^{t}\mathrm{d}s\;C(t',\,s)\frac{\lambda(t)}{\lambda(s)} = \sigma^2\int_{0}^{t'}\mathrm{d}s\;\lambda(t-r)\lambda(t'-r) - \frac{\lambda(t)}{\lambda(t')}\nu(t'). \tag{C.10}$$

Equation (C.8) depends explicitly on our previous solution for $W_1(x,\,t)$, equation (24). To obtain the desired solution we have to perform the integral in ω and inverse Fourier transform. This is most easily done if we first Fourier transform equation (C.8) in the first spatial variable, x, to give

$$\widehat{W_2}(k,\,t;\,k',\,t') = \int_{-\infty}^{\infty} \mathrm{d}\omega\; W_1(\omega,\,t')\mathrm{e}^{-\mathrm{i}\omega\left[k\frac{\lambda(t)}{\lambda(t')}+k'\right]-\mathrm{i}k\,[\Psi_d(t)-\Psi_d(t')]-kk'\beta(t',t)-k^2\rho(t',t)}. \tag{C.11}$$

One can now substitute equation (24) into this expression and performing the integral over ω, to obtain

$$\begin{aligned}\widehat{W_2}(k,\,t;\,k',\,t') = \int_{-\infty}^{\infty} \mathrm{d}z\; W_0(z)\exp\{&-\mathrm{i}k'[z\lambda(t') + \Psi_d(t')] \\ &- \mathrm{i}k\,[z\lambda(t) + \Psi_d(t)] - k'^2\frac{\nu(t')}{2} \\ &- k^2\left[\frac{\nu(t')\lambda^2(t)}{2\lambda^2(t')} + \rho(t',\,t)\right] - kk'\left[\beta(t',\,t) + \frac{\nu(t')\lambda(t)}{\lambda(t')}\right]\Bigg\}.\end{aligned} \tag{C.12}$$

Using equations (C.9) and (C.10), we can respectively evaluate the k^2 and the kk' terms in the above expression to be

$$\frac{\nu(t')\lambda^2(t)}{2\lambda^2(t')} + \rho(t',\,t) = \frac{\nu(t)}{2}\beta(t',\,t) + \frac{\nu(t')\lambda(t)}{\lambda(t')} = \sigma^2\int_0^{t'} \mathrm{d}s\;\lambda(t-s)\lambda(t'-s). \tag{C.13}$$

Substituting in these expressions, we see that equation (C.12) is a bivariate Gaussian distribution in the Fourier domain. Performing the double inverse Fourier transform we obtain the desired bivariate Gaussian distribution in the spatial domain, equation (25).

References

[1] Smoluchowski M 1916 Über Brownsche molekularbewegung unter einwirkung äußerer kräfte und deren zusammenhang mit der verallgemeinerten diffusionsgleichung *Ann. Phys. Leipzig* **353** 1103–12

[2] Chandrasekhar S 1943 Stochastic problems in physics and astronomy *Rev. Mod. Phys.* **15** 1–89

[3] Chandrasekhar S 1989 Marian Smoluchowski as the founder of the physics of stochastic phenomena, *Selected Papers vol 3: Stochastic, Statistical, and Hydromagnetic Problems in Physics and Astronomy* (Chicago, IL: University of Chicago Press) pp 275–82

[4] Fulinski A 1998 On marian Smoluchowski's life and contribution to physics *Acta Phys. Pol.* B **29** 1523–37

[5] Chandrasekhar S, Kac M and Smoluchowski R 2000 *Marian Smoluchowski: His Life and Scientific Work* (Warsaw: Polish Scientific Publishers)

[6] Piasecki J 2007 Centenary of marian Smoluchowski's theory of Brownian motion *Acta Phys. Pol.* B **38** 1623–9

[7] Metzler R, Jeon J-H, Cherstvy A G and Barkai E 2014 Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking *Phys. Chem. Chem. Phys.* **16** 24128–64

[8] Biswas J and Veloso M 2014 Episodic non-Markov localization: reasoning about short-term and long-term features *2014 IEEE Int. Conf. on Robotics and Automation (ICRA)* (Piscataway, NJ: IEEE) pp 3969–74

[9] Saavedra S, Hagerty K and Uzzi B 2011 Synchronicity, instant messaging, and performance among financial traders *Proc. Natl Acad. Sci. USA* **108** 5296–301

[10] Kyrychko Y N and Hogan S J 2010 On the use of delay equations in engineering applications *J. Vib. Control* **16** 943–60

[11] Erneux T 2009 *Applied Delay Differential equations* vol 3 (Berlin: Springer)

[12] Sipahi R, Vyhlídal T, Niculescu S and Pepe P 2012 *Time Delay Systems: Methods, Applications and New Trends* vol 423 (Berlin: Springer)

[13] Li X *et al* 2006 Discovery of time-delayed gene regulatory networks based on temporal gene expression profiling *BMC Bioinform.* **7** 26

[14] Zoppoli P, Morganella S and Ceccarelli M 2010 Timedelay-ARACNE: reverse engineering of gene networks from time-course data by an information theoretic approach *BMC Bioinform.* **11** 154
[15] Jaśkowski P 1982 Distribution of the human reaction time measurements *Acta Neurobiol. Exp.* **43** 221–5
[16] Nagy M, Ákos Z, Biro D and Vicsek T 2010 Hierarchical group dynamics in pigeon flocks *Nature* **464** 890–3
[17] Nagy M, Vásárhelyi G, Pettit B, Roberts-Mariani I, Vicsek T and Biro D 2013 Context-dependent hierarchies in pigeons *Proc. Natl Acad. Sci. USA* **110** 13049–54
[18] Giuggioli L, McKetterick T and Holderied M 2015 Delayed response and biosonar perception explain movement coordination in trawling bats *PLoS Comput. Biol.* **11** e1004089
[19] Voss H U and Kurths J 2002 Analysis of economic delayed-feedback dynamics *Modelling and Forecasting Financial Data* (*Studies in Computational Finance* vol 2) ed A Soofi and L Cao (Norwell, MA: Kluwer Academic Publishers) pp 327–49
[20] Lainscsek C and Sejnowski T J 2013 Electrocardiogram classification using delay differential equations *Chaos* **23** 023132
[21] Goldenfeld N 1992 *Lectures on Phase Transitions and the Renormalization Group* (*Frontiers in Physics* vol 85) (Reading, MA: Westview Press)
[22] Van Kampen N G 2007 *Stochastic Processes in Physics and Chemistry* 3rd edn (Amsterdam: Elsevier)
[23] Fox R F 1977 The generalized Langevin equation with gaussian fluctuations *J. Math. Phys.* **18** 2331–5
[24] Fox R F 1978 Gaussian stochastic processes in physics *Phy. Rep.* **48** 179–283
[25] Adelman S A 1976 Fokker–Planck equations for simple non-Markovian systems *J. Chem. Phys.* **64** 124–30
[26] Fox R F 1977 Analysis of nonstationary, Gaussian and non-Gaussian, generalized Langevin equations using methods of multiplicative stochastic processes *J. Stat. Phys.* **16** 259–79
[27] Hänggi P 1978 Correlation functions and master equations of generalized (non-Markovian) Langevin equations *Z. Phys.* **31** 407–16
[28] Sancho J M, San Miguel M, Katz S L and Gunton J D 1982 Analytical and numerical studies of multiplicative noise *Phys. Rev.* A **26** 1589
[29] Hernández-Machado A, Sancho J M, San Miguel M and Pesquera L 1983 Joint probability distribution of nonMarkovian sde *Z. Phys.* **52** 335–43
[30] Lutz E 2001 Fractional Langevin equation *Phys. Rev.* E **64** 051106
[31] Pottier N 2003 Aging properties of anomalously diffusing particule *Physica* A **317** 371–82
[32] Burov S and Barkai E 2008 Critical exponent of the fractional Langevin equation *Phys. Rev. Lett.* **100** 070601
[33] Hänggi P and Talkner P 1978 On the equivalence of time-convolutionless master equations and generalized Langevin equations *Phys. Lett.* A **68** 9
[34] San Miguel M and Sancho J M 1980 A colored-noise approach to Brownian motion in position space. Corrections to the Smoluchowski equation *J. Stat. Phys.* **22** 605–24
[35] Hänggi P and Jung P 1995 Colored noise in dynamical systems *Advances in Chemical Physics* vol 89 ed I Prigogine and S Rice (New York: Wiley) pp 239–326
[36] Guillouzic S, L'Heureux I and Longtin A 1999 Small delay approximation of stochastic delay differential equations *Phys. Rev.* E **59** 3970–82
[37] Ohira T and Milton J G 1995 Delayed random walks *Phys. Rev.* E **52** 3277
[38] Frank T D 2007 Kramers–Moyal expansion for stochastic differential equations with single and multiple delays: applications to financial physics and neurophysics *Phys. Lett.* A **360** 552–62
[39] Frank T D and Beek P J 2001 Stationary solutions of linear stochastic delay differential equations: applications to biological systems *Phys. Rev.* E **64** 021917
[40] Frank T D 2003 A Markov approach to nonlinear multivariate delay systems with noise *Phys. Scr.* **68** 333
[41] Longtin A 2010 Stochastic delay-differential equations *Complex Time-Delay Systems* (*Understanding Complex Systems*) ed F M Atay (Berlin: Springer) pp 177–95
[42] McKetterick T J and Giuggioli L 2014 Exact dynamics of stochastic linear delayed systems: application to spatiotemporal coordination of comoving agents *Phys. Rev.* E **90** 042135
[43] Chase M, McKetterick T J, Giuggioli L and Kenkre V M 2016 Langevin analysis of time-nonlocal Brownian motion with algebraic memories and delay interactions. *Eur. Phys. J.* B **87** 1–5

[44] Hutchinson J 1997 Relaxation processes and physical aging *The Physics of Glassy Polymers* (Berlin: Springer) pp 85–153
[45] Wang M C and Uhlenbeck G E 1945 On the theory of the Brownian motion: II *Rev. Mod. Phys.* **17** 323–42
[46] Bayes B T 1763 An essay towards solving a problem in the doctrine of chances *Phil. Trans.* **53** 370–418
[47] Hanggi P, Thomas H, Grabert H and Talkner P 1978 Note on time evolution of non-Markov processes *J. Stat. Phys.* **18** 155–9
[48] Hänggi P and Thomas H 1982 Stochastic processes: time evolution, symmetries and linear response *Phys. Rep.* **88** 207–319
[49] Isserlis L 1918 On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables *Biometrika* **32** 134–9
[50] Budini A A and Cáceres M O 2004 Functional characterization of linear delay Langevin equations *Phys. Rev.* E **70** 046104
[51] Bellman R and Cooke K L 1963 *Differential-Difference Equations* (Santa Monica, CA: RAND Corporation)
[52] Frank T D 2005 Delay Fokker–Planck equations, Novikov's theorem, and Boltzmann distributions as small delay approximations *Phys. Rev.* E **72** 011112
[53] Novikov E A 1965 Functionals and the random-force method in turbulence theory *Sov. Phys.—JETP* **20** 1290–4
[54] Hänggi P 1985 The functional derivative and its use in the description of noisy dynamical systems *Stochastic Processes Applied to Physics* ed L Pesquera and M A Rodriguez (Singapore: World Scientific) p 69
[55] Hänggi P 1989 Colored noise in continuous dynamical systems: a functional calculus approach *Noise in Non-Linear Dynamical Systems* ed F Moss and P V E McClintock (Cambridge: Cambridge University Press) pp 307–28
[56] Hänggi P and Thomas H 1977 Time evolution, correlations, and linear response of non-Markov processes *Z. Phys.* **26** 85–92
[57] Küchler U and Mensch B 1992 Langevins stochastic differential equation extended by a time-delayed term *Stoch. Stoch. Rep.* **40** 23–42
[58] Kenkre V M 2007 Analytic formulation, exact solutions, and generalizations of the elephant and the alzheimer random walks arXiv:0708.0034
[59] Chapman S 1916 On the law of distribution of molecular velocities, and on the theory of viscosity and thermal conduction, in a non-uniform simple monatomic gas *Phil. Trans. R. Soc.* A **216** 279–348
[60] Kolmogorov A N 1931 On analytical methods in the theory of probability *Math. Ann.* **104** 415–58
[61] McCauley J 2012 Non-Markov stochastic processes satisfying equations usually associated with a Markov process *Eur. Phys. J. Spec. Top.* **204** 133–43
[62] Mittag-Leffler G 1920 Sur la representation analytique d'une branche uniforme d'une fonction monogene *Acta Math.* **42** 285–328
[63] Gorenflo R, Kilbas A, Mainardi F and Rogosin S 2014 *Mittag-Leffler Functions, Related Topics and Application* (Berlin: Springer)
[64] Corless R M, Gonnet G H, Hare D E G, Jeffrey D J and Knuth D E 1996 On the Lambert *W* function *Adv. Comput. Math.* **5** 329–59
[65] Dechant A, Lutz E, Kessler D A and Barkai E 2014 Scaling Green–Kubo relation and application to three aging systems *Phys. Rev.* X **4** 011022
[66] Kalecki M 1935 A macrodynamic theory of business cycles *Econometrica* **3** 327–44
[67] Frisch R and Holme H 1935 The characteristic solutions of a mixed difference and differential equation occuring in economic dynamics *Econometrica* **3** 225–39
[68] Hunt D, Korniss G and Szymanski B 2010 Network synchronization in a noisy environment with time delays: fundamental limits and trade-offs *Phys. Rev. Lett.* **105** 068701
[69] Hod S 2010 Analytic treatment of the network synchronization problem with time delays *Phys. Rev. Lett.* **105** 208701
[70] Hunt D, Korniss G and Szymanski K 2011 The impact of competing time delays in coupled stochastic systems *Phys. Lett.* A **375** 880–5
[71] Wang Q, Chen G and Perc M 2011 Synchronous bursts on scale-free neuronal networks with attractive and repulsive coupling *PLoS One* **6** e15851

[72] Hunt D, Szymanski B K and Korniss G 2012 Network coordination and synchronization in a noisy environment with time delays *Phys. Rev.* E **86** 056114
[73] Hunt D, Molnár F Jr, Szymanski B and Korniss G 2015 Extreme fluctuations in stochastic network coordination with time delays *Phys. Rev.* E **92** 062816
[74] Ricciardi L M and Sacerdote L 1987 On the probability densities of an Ornstein–Uhlenbeck process with a reflecting boundary *J. App. Probab.* **24** 355–69
[75] Spendier K, Sugaya S and Kenkre V M 2013 Reaction-diffusion theory in the presence of an attractive harmonic potential *Phys. Rev.* E **88** 062142
[76] Guillouzic S 2000 Fokker–Planck approach to stochastic delay differential equations *PhD Thesis* Ottawa-Carleton Institute for Physics, University of Ottawa
[77] Frank T D, Beek P J and Friedrich R 2003 Fokker–Planck perspective on stochastic delay systems: exact solutions and data analysis of biological systems *Phys. Rev.* E **68** 021912
[78] Risken H 1984 *The Fokker–Planck Equation* (*Springer Series in Synergetics* vol 18) (Berlin: Springer)
[79] Shtelen W M and Stogny V I 1989 Symmetry properties of one-and two-dimensional Fokker–Planck equations *J. Phys. A: Math. Gen.* **22** L539