Mean, covariance, and effective dimension of stochastic distributed delay dynamics

Alexandre René, and André Longtin

Citation: Chaos **27**, 114322 (2017);
View online: https://doi.org/10.1063/1.5007866
View Table of Contents: http://aip.scitation.org/toc/cha/27/11
Published by the American Institute of Physics

Mean, covariance, and effective dimension of stochastic distributed delay dynamics

Alexandre René[1,a)] and André Longtin[1,2,b)]
[1]Department of Physics, University of Ottawa, Ottawa K1N 6N5, Canada
[2]Brain and Mind Research Institute, University of Ottawa, Ottawa K1H 8M5, Canada

(Received 6 May 2017; accepted 4 September 2017; published online 26 October 2017)

Dynamical models are often required to incorporate both delays and noise. However, the inherently infinite-dimensional nature of delay equations makes formal solutions to stochastic delay differential equations (SDDEs) challenging. Here, we present an approach, similar in spirit to the analysis of functional differential equations, but based on finite-dimensional matrix operators. This results in a method for obtaining both transient and stationary solutions that is directly amenable to computation, and applicable to first order differential systems with either discrete or distributed delays. With fewer assumptions on the system's parameters than other current solution methods and no need to be near a bifurcation, we decompose the solution to a linear SDDE with arbitrary distributed delays into natural modes, in effect the eigenfunctions of the differential operator, and show that relatively few modes can suffice to approximate the probability density of solutions. Thus, we are led to conclude that noise makes these SDDEs effectively low dimensional, which opens the possibility of practical definitions of probability densities over their solution space. *Published by AIP Publishing*. https://doi.org/10.1063/1.5007866

In order to accurately capture the dynamics of natural phenomena, mathematical models often include both transmission delays and a form of randomness. This work develops a computationally practical approach to decompose linear stochastic delay differential equations (SDDEs) into natural modes. We show that at least for single and uniformly distributed delays, only a few of these modes suffice to describe a solution, which allows solutions to be described as low dimensional probability densities.

I. INTRODUCTION

Delay differential equations (DDEs) can succinctly capture spatio-temporal interactions within a dynamical system. However, their mathematical analysis is complicated by their intrinsically infinite dimensional nature, a difficulty still compounded when one adds stochastic external forcing. Since stochastic delayed differential equations (SDDEs) appear in phenomena as diverse as El Niño oscillations,[7,8] resonance in optical cavities,[15,21] cell and gene regulation,[6,25,27] and neural population dynamics,[2,23,34] there is a need to develop appropriate analytical methods for these types of equations.

While much of the contemporary interest in SDDEs stems from models of nonlinear phenomena, a common analytical approach is to linearize the models around points of interest.[3,7,21,23,25] Of particular interest are the moments of the linearized dynamical variable, as they frequently appear in expressions for the nonlinear system [e.g., Eq. (5) in Ref. 23]. In particular, the autocovariance is frequently used to quantify temporal structure.[6,7,13,27]

Since our focus is on linear systems, it is tempting to seek a solution in terms of orthogonal eigenvectors. The challenge then is that the dual eigenvectors, which together with the eigenvectors form a biorthogonal system, are difficult to find for a delayed system;[36] to our knowledge, they have only been derived for discrete delays.[1,5] Amann *et al.*,[1] for example, consider an equation for a real-valued process $X_0 : \mathbb{R} \to S \subseteq \mathbb{R}$ of the form

$$dX_0(t) = (a X_0(t) + b X_0(t-1))\, dt + q\, dW(t), \qquad (1)$$

where $a, b, q \in \mathbb{R}$, and dW is the differential of a Wiener process. They provide direct computational expressions for the dual eigenfunctions by first considering the case $q = 0$ and then treating the noise as a forcing term. Although their noise-free result is a special case of the decomposition of functional differential equations (FDEs), it is elegantly and succinctly derived with no appeal to FDE theory. To our knowledge, they were also the first to treat an SDDE by means of eigenfunction decomposition.

The results in this paper expand on the efforts of Amann *et al*. First, we develop a systematic approach to obtain the biorthogonal decomposition. We use this to generalize the results of Amann *et al*. to the more general distributed delay case

$$dX_0(t) = \left(\int_0^{\tau_{max}} \kappa(\tau) X_0(t-\tau)\, d\tau \right) dt + q\, dW(t), \qquad (2)$$

where X_0 is again a real-valued process, $0 < \tau_{max} < \infty$, q and dW are as above, and both κ and the initial condition are integrable functions on $[0, \tau_{max}]$. Equation (1) can be recovered from Eq. (2) by setting $\kappa(\tau) = a\delta(\tau) + b\delta(\tau - 1)$,

a)Electronic mail: arene010@uottawa.ca.
b)Electronic mail: alongtin@uottawa.ca; URL: http://mysite.science.uottawa.ca/alongtin/

where δ is the Dirac delta function. A sufficient condition for our approach is that the operator defined by kernel κ be non-degenerate. We also relate our expressions to known results from the theory of FDEs, providing a more robust context for further extensions.

Having established the tools for an eigenfunction decomposition, we then write the full time-dependent solution for the moments of Eq. (2) as series over eigenvalues; this differs from other efforts[1,14] where only the stationary solution is considered. We show that in most cases of interest, this series may be safely truncated to a relatively small number of terms.

Solutions to an SDDE are often given as a time-dependent probability density function (p.d.f.) over S. However, just as fully defining the state of a DDE requires a function over the interval $(0, \tau]$, for an SDDE we should seek a p.d.f. over a space of functions from $(0, \tau]$ to S. As this space is infinite-dimensional, such p.d.f.'s are difficult to estimate from experimental or numerical data. Our eigenfunction decomposition offers a potential solution to this problem: Since only a limited number of modes suffice to describe a state, we can marginalize the p.d.f. onto these modes. We describe this as the p.d.f. being "effectively finite-dimensional."

In addition to being applicable to distributed delays, our approach places very little restriction on either the feedback kernel κ or the noise amplitude q. This is in contrast to other current approaches based on either solving the delay Fokker-Planck equation[13,14,16] or reducing the system to its center manifold:[19,24] Fokker-Planck solutions have at present only been developed for single delayed systems such as Eq. (1), with small delayed feedback b and either short or very long delay τ. On the other hand, while a centre manifold reduction also provides a low-dimensional representation similar to our own, it can only be performed in the vicinity of certain bifurcations. Moreover, the dimensionality of this representation is fixed by that of the centre manifold, whereas in the eigendecomposition, it can be adjusted to obtain arbitrary accuracy.

Thus, two advantages of our approach are its general applicability to the entire parameter space and the adaptability of its low-dimensional representation. In addition, because it is developed entirely in terms of limits of finite-dimensional matrix operators, the expressions we provide are easily computable and can be obtained to arbitrary precision using standard linear algebra numerical routines. Although this requires mild additional assumptions compared to the functional analysis approach, it allows for a presentation that is both more intuitive and closely parallels how SDDEs are solved numerically.

For clarity and pedagogical reasons, our treatment of Eq. (2) will be limited to cases where X_0 is scalar. Just as the FDE theory can be formulated for multivariate equations, so too can our theory be extended to the multivariate case. Moreover, the overlap between our results and known results for FDEs is then preserved.

This paper is organized as follows. In Sec. II, we begin by reviewing some key results of FDEs as they relate to delay equations. We then extend these to stochastic equations in Sec. III and obtain the solution to Eq. (2). Section IV follows with a derivation of the time-dependent autocovariance function for our problem of interest. Section V examines the issue of series truncation, deriving an expression for each eigenfunction's relative contribution and discussing its implications with respect to the dimensionality of solutions. Finally, we conclude with an outlook of possible future directions.

II. EIGENFUNCTION DECOMPOSITION OF FDEs

A delay equation may be cast as a retarded functional differential equation (RFDE) on the set of continuous functions $C \equiv C([0, \tau_{max}], \mathbb{R})$; see, for example, Ref. 18 (Secs. 1.2 and 7). The state $x(t)$ is then viewed as a continuous function

$$\begin{aligned} x(t) : [0, \tau_{max}] &\to \mathbb{R} \\ \tau &\mapsto x_\tau(t), \end{aligned} \tag{3}$$

for which

$$x_\tau(t) \stackrel{\text{def}}{=} x_0(t-\tau), \quad \tau \in [0, \tau_{max}]\,. \tag{4}$$

This notation helps suggest the interpretation of t as some continuous index of an infinite dimensional vector $x(t)$; it also helps distinguish $x_t(t) \in \mathbb{R}$ from $x(t) \in C$.

We note here that following the usual convention of reserving uppercase letters for random variables, the state x in Eq. (4) is denoted in lowercase since RFDEs are deterministic.

To treat the noise-free case of Eq. (2) as an RFDE, we recast the scalar variable x_0 as a function $x \in C$, which evolves over time. The right endpoint $x_0(t)$ tracks our variable of interest; other points satisfy Eq. (4) and keep track of the variable's history. Then, Eq. (2), with $q = 0$, may be rewritten as (see Ref. 18, p. 194)

$$\frac{dx(t)}{dt} = \boldsymbol{A}x(t) \in C, \tag{5}$$

where the function $\boldsymbol{A}x(t)$, taking values for $\tau \in [0, \tau_{max}]$, is given by

$$[\boldsymbol{A}x(t)]_\tau = \begin{cases} \displaystyle\int_0^{\tau_{max}} \kappa(\nu)x(t-\nu)\,d\nu & \tau = 0 \\ \displaystyle -\frac{dx_\nu(t)}{d\nu}\Big|_{\nu=\tau} & 0 < \tau \le \tau_{max}. \end{cases}$$

The eigenfunctions of $\boldsymbol{A}$ take the form

$$\phi_{n,\tau} = e^{-\lambda_n \tau}, \quad \tau \in [0, \tau_{max}], \tag{6}$$

where λ_n is a solution to (Ref. 18, p. 198)

$$\lambda_n - \int_0^{\tau_{max}} \kappa(\tau) e^{-\lambda_n \tau}\,d\tau = 0\,. \tag{7}$$

Following the same convention as for x, $\phi_n \in C$ here denotes the eigenfunction associated with eigenvalue λ_n, while $\phi_{n,\tau} \in \mathbb{R}$ denotes the evaluation of that eigenfunction at a particular lag τ.

It is known that Eq. (7) has at most finitely many solutions with positive real part (Ref. 18, Sec. 7.6). Moreover, under some mild conditions,[28] solutions with negative real part are infinitely many and the eigenfunctions ϕ_n form a basis that spans the entire space C. (In some cases, such as when κ is a gamma function, the operator $\boldsymbol{A}$ may be degenerate and have only a finite number of eigenvalues. One must then extend the basis to generalized eigenfunctions.) Solutions to Eq. (2) can therefore be written as

$$x(t) = \sum_{\lambda_n} c_n e^{\lambda_n t}\phi_n, \quad c_n \in \mathbb{C}, \tag{8a}$$

in the FDE view, or equivalently

$$x_0(t) = \sum_{\lambda_n} c_n e^{\lambda_n t}\phi_{n,0}, \quad c_n \in \mathbb{C}, \tag{8b}$$

in the DDE view. To obtain the c_n, we need the functionals $\phi_n^\dagger$, dual to ϕ_n, such that

$$\phi_n^\dagger(\phi_m) = \delta_{nm} \equiv \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise;} \end{cases} \tag{9}$$

these are given in Ref. 18, p. 211 as:

$$\phi_n^\dagger(x) = \mathcal{N}_n x_0 + \mathcal{N}_n \int_0^{\tau_{max}} d\tau \int_\tau^{\tau_{max}} d\nu\, \kappa(\nu) e^{\lambda_n(\tau-\nu)} x_\tau, \tag{10}$$

with the normalization constant $\mathcal{N}_n$ chosen to satisfy Eq. (9). Some care must be taken when evaluating $\mathcal{N}_n$ for kernels κ with a Dirac delta at 0, which is why the implicit definition through Eq. (9) is often preferred. We will give an explicit expression in Sec. III.

It has been argued[1] that the practical applicability of the expansion in Eq. (8) is limited by the fact that there are an infinite number of eigenvalues. However, we will show in Sec. V that the eigenfunctions are in fact strongly hierarchized, and thus, that the series in Eq. (8) can be truncated to a given desired accuracy.

III. EIGENFUNCTION DECOMPOSITION OF SDDEs

A standard approach in the study of especially linear DDEs is to discretize the differential operator in time.[4,20,31,33] Here, we extend this method to the linear SDDE of Eq. (2). Doing so allows us to recast the problem as a finite-dimensional multivariate SDE, for which solutions are already known. We then obtain continuous time solutions by making the discretization infinitely fine. The discretization is thus only an intermediate step, useful for defining the stochastic equation but ultimately absent from the final expressions.

As in Sec. II, we make the assumption that the eigenfunctions of $\boldsymbol{A}$ span C. In addition, our derivation of the results in Sec. IV also assumes $\boldsymbol{A}$ to have only finitely many real eigenvalues.

When discretized into $N-1$ time steps, $\boldsymbol{A}$ becomes the $N \times N$ dimensional matrix $\boldsymbol{A}^{(N)}$ defined in Eq. (13) (see Appendix B for details). In this form, the vector $\boldsymbol{x}^{(N)}(t)$ corresponds to the function $x(t)$ sampled at regular intervals $\frac{\tau_{max}}{N}$. Equation (5) is then approximated by the linear ODE

$$\frac{d\boldsymbol{x}^{(N)}}{dt} = \boldsymbol{A}^{(N)}\boldsymbol{x}^{(N)}, \tag{11}$$

for which eigenvalues $\lambda_n^{(N)}$ and eigenvectors $\boldsymbol{\phi}_n^{(N)}$, $\boldsymbol{\phi}_n^{(N)\dagger}$ can be computed. In Appendix B, we show that these are equal to the expressions for λ_n, ϕ_n, and $\phi_n^\dagger$ given in Sec. II, in the limit $N \to \infty$.

To extend this discretization process to stochastic equations, we define $\boldsymbol{Q}^{(N)} \in \mathbb{R}^N$,

$$Q_n^{(N)} \equiv q\,\delta_{n0}, \quad q \in \mathbb{R}, \tag{12}$$

where δ_{n0} is the Kronecker delta defined in Eq. (9). Equation (2) can then be approximated by the multivariate SDE

$$d\boldsymbol{X}^{(N)} = \boldsymbol{A}^{(N)}\boldsymbol{X}^{(N)}dt + \boldsymbol{Q}^{(N)}dW, \tag{13a}$$

which expands to

$$\begin{pmatrix} dX_0 \\ dX_1 \\ \vdots \\ dX_{N-1} \end{pmatrix} \approx \underbrace{\begin{pmatrix} \bar\kappa(0)\frac{\tau_{max}}{N} & \bar\kappa\left(\frac{\tau_{max}}{N}\right)\frac{\tau_{max}}{N} & & \bar\kappa\left(\frac{(N-1)}{N}\tau_{max}\right)\frac{\tau_{max}}{N} \\ N/\tau_{max} & -N/\tau_{max} & & \\ 0 & N/\tau_{max} & -N/\tau_{max} & \\ & & \ddots & \\ & & N/\tau_{max} & -N/\tau_{max} \end{pmatrix}}_{\equiv \boldsymbol{A}^{(N)}} \begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_{N-1} \end{pmatrix} dt + \underbrace{\begin{pmatrix} q \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{\equiv \boldsymbol{Q}^{(N)}} dW, \tag{13b}$$

where

$$\bar\kappa(t) = \frac{1}{\Delta t}\int_t^{(t+\Delta t)^-} \kappa(s)\,ds, \quad \Delta t = \frac{\tau_{max}}{N}.$$

The discretized kernel $\bar\kappa$ is defined such that if κ is given by a Dirac delta function, i.e., $\kappa(\tau) = \delta(t - n\Delta t)$, then $\bar\kappa(n\Delta t) = 1$ and $\bar\kappa((n+1)\Delta t) = 0$.

Since Eq. (13) describes a multivariate Ornstein-Uhlenbeck process, its solution is a multivariate Gaussian of the form

$$\boldsymbol{X}^{(N)}(t) \sim \mathcal{N}(\boldsymbol{\mu}^{(N)}(t), \boldsymbol{\Sigma}^{(N)}(t)). \tag{14}$$

Assuming that $\boldsymbol{A}^{(N)}$ is full rank and that the initial condition $x^{(N)}(0)$ is known, we can express the mean vector $\boldsymbol{\mu}^{(N)}$ and

covariance matrix $\mathbf{\Sigma}^{(N)}$ in terms of eigenvectors [see, e.g., Ref. 32, Eqs. (6.122) and (6.123) or Ref. 30, Eqs. (2.54)–(2.56)]. For the mean we have

$$\boldsymbol{\mu}^{(N)}(t) = \sum_{\lambda_n} c_n(t)\boldsymbol{\phi}_n^{(N)}, \tag{15a}$$

where

$$c_n(t) = e^{\lambda_n^{(N)} t}\boldsymbol{\phi}_n^{(N)\dagger} x^{(N)}(0), \tag{15b}$$

while the covariance is given by

$$\Sigma^{(N)}(t) = \sum_{n,m} D_{n,m}\frac{e^{(\lambda_n^{(N)}+\lambda_m^{(N)})t}-1}{\lambda_n^{(N)}+\lambda_m^{(N)}}\boldsymbol{\phi}_n^{(N)}\left(\boldsymbol{\phi}_m^{(N)}\right)^{\mathsf{T}}, \tag{16a}$$

with

$$D_{n,m} = \boldsymbol{\phi}_n^{(N)\dagger} Q_n^{(N)} Q_n^{(N)\mathsf{T}}\left(\boldsymbol{\phi}_m^{(N)\dagger}\right)^{\mathsf{T}}. \tag{16b}$$

We note that $\boldsymbol{\phi}^{(N)\mathsf{T}}$ and $\boldsymbol{\phi}^{(N)\dagger}$ are not equivalent. The first indicates usual vector transposition of $\phi^{(N)}$, while the latter denotes the member of the dual associated with $\phi^{(N)}$ as given by Eq. (B7c). Thus $\phi^{(N)}$ and $(\boldsymbol{\phi}^{(N)\dagger})^{\mathsf{T}}$ are both column vectors, but distinct. Due to the particular form of $Q_n^{(N)}$ [Eq. (12)], the covariance matrix only depends on the first component of each dual eigenvector.

To obtain an exact solution to Eq. (2), we map each discretized step k to its corresponding time τ

$$\tau \equiv k/N, \quad k \in \{0, 1, \ldots, N-1\}, \tag{17}$$

and let the number of steps N go to infinity. As shown in Appendix B, in this case eigenvectors become eigenfunctions

$$\phi_{n,t} = e^{-\lambda_n \tau}, \tag{18a}$$

consistent with Sec. II. The full solution can be written

$$X(t) \sim \mathcal{N}(\mu(t), \Sigma(t)), \tag{18b}$$

$$\mu(t) = \sum_n e^{\lambda_n(t-\tau)}\phi_n^\dagger(x(0)), \tag{18c}$$

$$\Sigma_{\tau\nu}(t) = q^2\sum_{n,m}\mathcal{N}_n\mathcal{N}_m\frac{e^{(\lambda_n+\lambda_m)t}-1}{\lambda_n+\lambda_m}e^{-\lambda_n\tau}e^{-\lambda_m\nu}, \tag{18d}$$

where

$$\phi_n^\dagger x = \mathcal{N}_n x_0 + \mathcal{N}_n\int_{0^+}^{\tau_{max}} dt\int_t^{\tau_{max}} d\nu\,\kappa(\nu)e^{\lambda_n(\tau-\nu)}x_\tau, \tag{18e}$$

$$\mathcal{N}_n = \left[1 + \int_{0^+}^{\tau_{max}}\kappa(\tau)\tau e^{-\lambda_n\tau}\,d\tau\right]^{-1}. \tag{18f}$$

Equations (18b)–(18d) are analogous to Eqs. (14)–(16), the only difference being the replacement of discretized quantities by their continuous limit. Equations (18a), (18e), and (18f) are similarly obtained from their discrete analogues, which are derived in Appendix B. We note that the lower bound of 0^+ in Eq. (18f), which permits the correct treatment of kernels κ containing a delta function at $\tau = 0$, is obtained naturally from the limit of the discrete equation.

We also note that Eq. (18e) takes the same value whether the lower bound is set to 0 or 0^+, as the integral over τ removes any contribution from a Dirac delta at $\nu = 0$. Equation (18e) is therefore equivalent to Eq. (10); we use 0^+ as the lower bound here to better reflect the result of our discretization procedure. This form also has the advantage of making the exclusion of $\kappa(0)$ from the integral explicit.

In contrast to Ref. 1, the series in Eqs. (18) is absolutely convergent for $\tau + \nu < \tau_{max}$ (a property we prove in Appendix E). Consequently, while in Ref. 1 it is necessary to define a non-trivial "sum rule," here the order of the terms can be chosen for convenience, as it does not impact the value of the series.

Although Eq. (18c) is formally equivalent to Eq. (8a), we must take care in our interpretation of the variable τ. Since it is defined by Eq. (17), τ in Eqs. (18) should be restricted to the interval

$$\tau \in I_{\text{lag}} \equiv [0, \tau_{max}). \tag{19}$$

We note that the exclusion of τ_{max} from I_{lag} is necessary, as evidenced by Fig. 1. Within this interval τ may be interpreted as a lag.

Equations (18) describe the full functional solution to Eq. (2). In the SDDE context, one is often more interested in

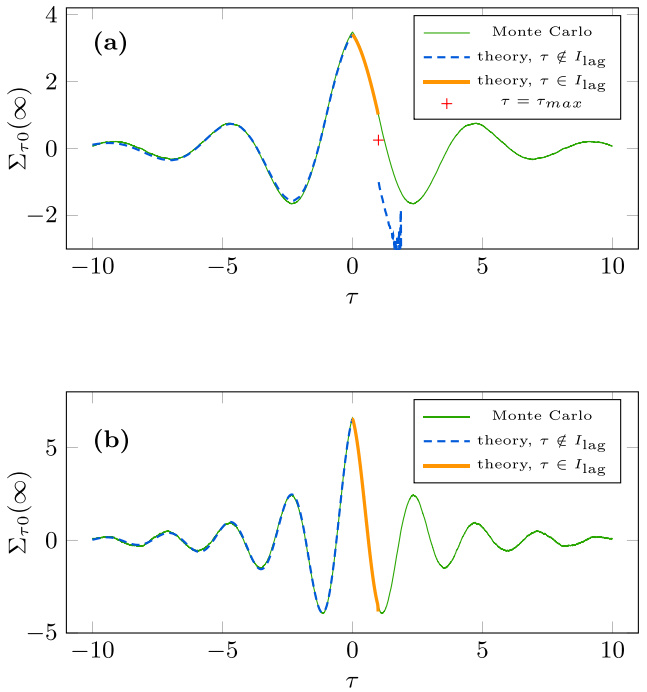

FIG. 1. (a) Comparison of Eq. (25) to a Monte Carlo estimate of the autocovariance of the stationary state, for Eq. (2) with a single delay $\kappa(\tau) = -\sqrt{2}\,\delta(\tau - 1)$. The Monte Carlo estimates are averaged over 10 000 realizations; Monte Carlo values for $\tau > 0$ are obtained by exploiting the autocovariance's symmetry. Equation (25) is expected to yield correct values for all lags τ within I_{lag}. It indeed starkly deviates from the estimate for $\tau \geq \tau_{max}$ (and diverges for $\tau \geq 2\,\tau_{max}$), but interestingly both curves remain in agreement for all negative time lags—despite the fact that our theory was derived only for $\tau \in I_{\text{lag}}$. (b) Same as (a), but with a constant delay kernel $\kappa(\tau) = -3$, $\tau \in [0, 1]$ and $q = 3$. For $\tau \geq \tau_{max}$, the theoretical curve exceeds the bounds of the figure and is not shown.

$X_0(t)$—the marginalization of $X(t)$ over all lags other than zero. This "pointwise" solution, which corresponds to averaging over all realizations and considering only the distribution of X_0 at time t, is easily obtained from Eqs. (18) as

$$X_0(t) \sim \mathcal{N}(\mu_0(t), \Sigma_{00}(t)) . \tag{20}$$

In contrast to Eq. (18b), Eq. (20) describes a one dimensional distribution, which allows us to verify the result numerically. We do this by comparing the time-dependent p.d.f. of Eq. (20) to time-slice histograms obtained by repeatedly integrating Eq. (2) with different noise realizations. The result is shown in Fig. 2 for both a single delay process where κ is proportional to $\delta(\tau - \tau_{max})$, and a constant delay process where κ is uniform over I_{lag}.

IV. EXTRACTING AUTOCOVARIANCE

While $\Sigma_{00}(t)$ describes the instantaneous variance of the process given by Eq. (2), the other components of the covariance matrix (18d) describe the temporal structure of the process. Indeed, since

$$\Sigma_{\tau\nu}(t) \overset{\text{def}}{=} \mathrm{E}[(X_\tau(t) - \mu_\tau(t))(X_\nu(t) - \mu_\nu(t))], \tag{21}$$

$\Sigma_{\tau\nu}(t)$ is precisely the autocovariance function. As such, it should also satisfy

$$\Sigma_{\tau\nu}(t) = \Sigma_{\nu\tau}(t), \tag{22}$$

$$\Sigma_{\tau+\sigma,\nu+\sigma}(t) = \Sigma_{\tau\nu}(t-\sigma), \tag{23}$$

where we have used the same subscript convention as in Eq. (4). Property (23) implies that there is some redundancy in the lag parameters, which we will remove later by setting ν to 0.

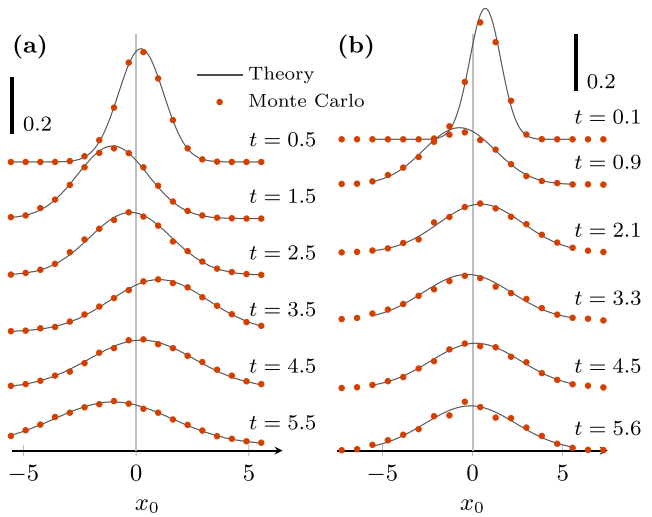

FIG. 2. Probability density function (p.d.f.) of $X_0(t)$ at different points in time. Dots show the approximate p.d.f. obtained by performing 10 000 realizations of the system given by Eq. (2); lines show the p.d.f. predicted by Eq. (18b). Both comparisons (a) and (b) take a known initial condition, $p(X_0(t) = x_0(t)) = \delta(x_0(t) - 1)$, $t \in [-\tau_{max}, 0]$. (a) Single delay kernel: $\kappa(\tau) = -1.56\,\delta(\tau - 1)$. Noise strength is set to $q = \sqrt{2}$. (b) Uniform delay kernel: $\kappa(\tau) = -3$, $\tau \in [0, 1]$. Noise strength is $q = 3$.

In practice, to evaluate the covariance matrix of Eq. (18d), the sum must be truncated, and so, convergence speed becomes important. Two factors influence this speed: the values of τ and ν, and the truncation scheme used. The necessity of a truncation scheme is due to the unspecified ordering of the double sum—for more details, see Appendix C. While a rigorous characterization of the convergence of Eq. (18d) is beyond the scope of this study, rearranging the terms in the order of decreasing importance suggests that convergence should be similar to that of an alternating p-series (Fig. 3). This is corroborated by an analysis of the absolute series' asymptotic behavior, described in Appendix E. The key result from this analysis is Eq. (E9), which can be summarized as

$$\begin{aligned}\Sigma_{\tau\nu}(t) &= \sum_{n,m} |\Phi_{\tau,\nu,t}(n,m)| \\ &\sim B'_{\tau,\nu,t} + B_{\tau,\nu} \sum_{k=n'}^{\infty} \frac{\log k}{k^{2-(\tau+\nu)/\tau_{max}}} \\ \Phi_{\tau,\nu,t}(n,m) &\equiv q^2 \mathcal{N}_n \mathcal{N}_m \frac{e^{(\lambda_n+\lambda_m)t} - 1}{\lambda_n + \lambda_m} e^{-\lambda_n \tau} e^{-\lambda_m \nu},\end{aligned} \tag{24}$$

where $B'_{\tau,\nu,t}$ and $B_{\tau,\nu}$ are constants. Since the $\log k$ becomes negligible compared to the denominator for large k, the r.h.s. of Eq. (24) indeed behaves as a p-series with $p = 2 - (\tau + \nu)/\tau_{max}$. Considering that Eq. (18d) is most likely comparable to an *alternating* p-series, we make the following supposition, which we expect to be true for any kernel κ such that the eigenfunctions of $\boldsymbol{A}$ span C:

Conjecture 1. *$\Sigma_{\tau\nu}(t)$, as defined in Eq. (18d), is absolutely convergent for $\tau + \nu < \tau_{max}$ and conditionally convergent for $\tau_{max} \leq \tau + \nu < 2\,\tau_{max}$.*

This validates Eq. (18d), as it guarantees that the series is convergent on the domain $\tau \in I_{\text{lag}}$ on which it was derived. Moreover, the lower the sum $\tau + \nu$, the faster we expect the convergence to be. Thus, we might as well fix the second lag ν to its minimum value of 0. This leads us to define the autocovariance function as

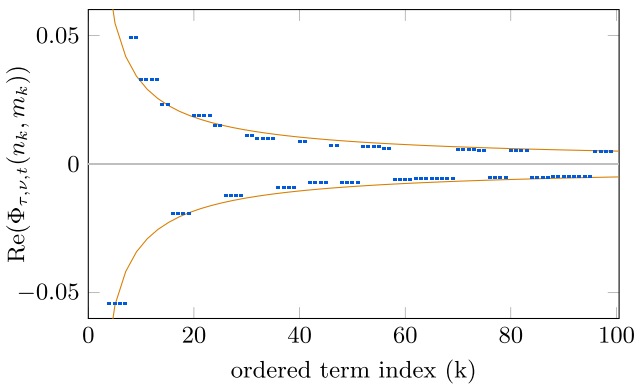

FIG. 3. When the terms of Eq. (24) are ordered such that $|\mathrm{Re}(\Phi_{\tau,\nu,t}(n_k, m_k))|$ is a monotonically decreasing function of k, with $\{n_k\}$ and $\{m_k\}$ reordered sequences of the indices n and m, they form a sequence roughly similar to an alternated p-series. Shown here is such a ordering for the terms of the single delay autocovariance with parameters $\tau = \nu = 0.6$ and $t = \infty$, where we have kept the 5th to 100th terms (the first four terms fall beyond the figure's boundaries). The solid lines are guides for the eye: They correspond to the functions $\pm 0.5\,k^{-0.8}$.

$$C_{XX}(t,s) \equiv \Sigma_{t-s,0}(t) = q^2 \sum_{n,m} \mathcal{N}_n \mathcal{N}_m \frac{e^{(\lambda_n+\lambda_m)t}-1}{\lambda_n+\lambda_m} e^{-\lambda_n(t-s)}, \tag{25}$$

where we have assumed, without loss of generality, that $t \geq s$. The case of the stationary state, i.e.,

$$\Sigma_{\tau 0}(\infty) = q^2 \sum_{n,m} \mathcal{N}_n \mathcal{N}_m \frac{-e^{-\lambda_n \tau}}{\lambda_n+\lambda_m} \tag{26}$$

was previously reported in Ref. 1, Eq. (27). We note that the autocorrelation function can be obtained from Eq. (25) by dividing by $\sqrt{\Sigma_{00}(t)\Sigma_{00}(s)}$.

It is important to remember that Eq. (25) was derived only for $t-s \in I_{\text{lag}} = [0, \tau_{max})$. Indeed, the comparison of Eq. (26) to Monte Carlo simulations shown in Fig. 1 clearly illustrates that $\Sigma_{\tau,0}$ is *not* in interpretable as the autocovariance for $\tau, \nu \geq \tau_{max}$. Figure 1 also suggests that we might be able to compute the autocovariance for arbitrarily long lag times by using negative lags, a possibility we leave for future investigations.

V. DIMENSIONALITY REDUCTION OF SDDEs

A. Amplitude equations

A decomposition onto eigenfunctions, or modes, gives us a corresponding description in terms of mode amplitudes. Thus, rather than studying the evolution of X subjected to the process given by Eq. (2), we can equivalently study the evolution of the projections $C_n(t) = \phi_n^\dagger(X(t))$ onto the modes ϕ_n. One has to be careful because $X(t)$ is a functional random variable, which makes the quantity $C_n(t)$ hard to define, but we can get around this by making use again of the discretized system of Eq. (13). Defining $C_n^{(N)}(t) \equiv \boldsymbol{\phi}_n^{(N)}(X^{(N)}(t))$, we can write its differential as

$$\begin{aligned} dC_n^{(N)}(t) &= d\left(\boldsymbol{\phi}_n^{(N)\dagger}(X^{(N)}(t))\right) = \boldsymbol{\phi}_n^{(N)\dagger}(dX^{(N)}(t)\Big) \\ &= \boldsymbol{\phi}_n^{(N)\dagger}\left(\boldsymbol{A}^{(N)}X^{(N)}(t)dt + \boldsymbol{Q}^{(N)}dW(t)\right) \\ &= \boldsymbol{\phi}_n^{(N)\dagger}\left(\sum \lambda_m^{(N)} C_m^{(N)}(t)\boldsymbol{\phi}_m^{(N)}\right)dt \\ &\quad + \underbrace{\boldsymbol{\phi}_n^{(N)\dagger}\boldsymbol{Q}^{(N)}}_{q\mathcal{N}_n^{(N)}} dW(t). \end{aligned}$$

This has a well defined limit as $N \to \infty$, and so we obtain

$$dC_n = \lambda_n C_n\, dt + q\mathcal{N}_n\, dW\,. \tag{27}$$

Regarding Eq. (27), we can immediately note that (i) it is not a delay equation; (ii) the amplitudes C_n are complex; and (iii) each C_n can be solved independently, even though the C_n are coupled through their shared noise.

From Eq. (27), for each mode C_n we can compute the mean amplitude

$$\langle C_n(t)\rangle = e^{\lambda_n t}\langle C_n(0)\rangle \tag{28}$$

and mean power

$$\langle |C_n(t)|^2\rangle = e^{2\text{Re}(\lambda_n)t}\left(\langle |C_n(0)|^2\rangle + \frac{q^2|\mathcal{N}_n|^2}{2\text{Re}(\lambda_n)}\right) - \frac{q^2|\mathcal{N}_n|^2}{2\text{Re}(\lambda_n)}; \tag{29}$$

with $\langle\cdot\rangle$ denoting the average over realizations. Appendix D details how we obtained these expressions. In particular, once the system reaches the stationary state (assuming it exists), the dependence on initial conditions disappears and we have

$$\langle |C_n(\infty)|^2\rangle = -\frac{q^2|\mathcal{N}_n|^2}{2\text{Re}(\lambda_n)}. \tag{30}$$

B. Relative importance of dynamical modes

Equations (29) and (30) indicate how much power each mode is expected to contribute. In particular, this contribution decreases monotonically with mode frequency [see Eq. (E6)]. This means that after some time t, the high-frequency components of the initial condition die off, leaving most of the power concentrated in relatively few components.

By quantifying this effect, we obtain an assessment of the *effective* modes, that is, those that are significantly contributing to the variance of a solution. We do this by computing each mode's relative power (RP):

$$RP(n,t) \equiv \frac{\langle |C_n(t)|^2\rangle}{\sum_{k=-\infty}^{\infty}\langle |C_k(t)|^2\rangle}. \tag{31}$$

We obtain the relative power of the stationary state by evaluating Eq. (31) at $t=\infty$:

$$RP(n,\infty) = \frac{|\mathcal{N}_n|^2/\text{Re}(\lambda_n)}{\sum_{k=-\infty}^{\infty}(|\mathcal{N}_k|^2/\text{Re}(\lambda_k))}. \tag{32}$$

In Table I, we compute Eq. (32) for the system with single delay τ

$$dX_0 = (a\,X_0 + b\,X_\tau)\,dt + q\,dW \tag{33}$$

for values of $\alpha \equiv a\tau$ and $\beta \equiv b\tau$. (Consistent with the notation introduced in Sec. II, X_τ here is the retarded solution, that is, $X_\tau(t) \overset{\text{def}}{=} X_0(t-\tau)$.) We see that very few modes—often only the first two conjugate ones—suffice to get a good approximation of this process' behavior. Progressively more are needed as one increases the strength of the instantaneous (i.e., non-delayed) term α. One also notes that at the stability boundaries, where some eigenvalues have zero real part, only the most dominant mode is present. This is expected, as in these cases the dynamics eventually collapse onto the centre manifold. In fact, the 7th and 14th lines in Table I correspond to Hopf bifurcations, for which the collapse to a two-dimensional system was already noted and exploited for centre manifold analysis.[19]

We are thus led to the interesting realization that noise *reduces* the effective dimensionality of an SDDE: While the deterministic component of the process attempts to bring the amplitude of each mode to zero following Eq. (28), the noisy component of the process simultaneously adds to each mode an amount of power proportional to q, following the ratio given by Eq. (31). As this ratio tends towards Eq. (32), it

TABLE I. Relative mode powers for the process (33) as computed by Eq. (32). Values of 0, 0.5 and 1 were obtained analytically while the others were computed numerically; all hold for any $q \neq 0$. Each mode corresponds to an eigenvalue; conjugate pairs lead to two modes of equal power. Parameter pairs were chosen to cover the range of expected stability conditions for this system; those for which $\alpha \geq -1$ are indicated on the phase diagram of Fig. 4 as colored $\times$ marks. Unstable systems have non-physical values for relative power (i.e., negative, or improperly normalized) since they do not have a stationary state. Systems on the stability boundary have eigenvalues with zero real part, which completely dominate. Systems with no delay at all have only a single mode.

		Relative stationary power		
α	β	Mode 1	Mode 2	Mode 3
1	-1	1	0	0
0.5	-1	0.499	0.499	0.001
0.5	-2	0.508	0.508	-0.006
0	-0.1	0.995	0.004	0.0003
0	-0.3	0.893	0.105	0.0006
0	-1	0.496	0.496	0.003
0	$-\pi/2$	0.5	0.5	0
-1	1	1	0	0
-1	0.5	0.952	0.019	0.019
-1	0	1	–	–
-1	-0.5	0.484	0.484	0.010
-1	-1	0.481	0.481	0.012
-1	-1.75	0.489	0.489	0.007
-1	-2.261	0.5	0.5	0
-5	-1	0.360	0.360	0.081
-10	-1	0.231	0.231	0.12
-100	99	0.152	0.127	0.127
-100	-1	0.023	0.023	0.023
-100	-100	0.410	0.410	0.045

privileges components associated with small $|n|$. Since those components come to dominate irrespective of the initial conditions, we can neglect the others, thereby reducing the effective dimensionality of the process.

This partial collapse of the process along certain dominant dimensions suggests that we might speak of a "dimensional hierarchy," in that although the system is strictly infinite dimensional, it can be approximated to arbitrary precision by a finite dimensional process. Such a low-dimensional representation is distinct from that obtained by reducing to the centre manifold, as it does not depend on the presence of a bifurcation and thus is valid over the entire parameter space.

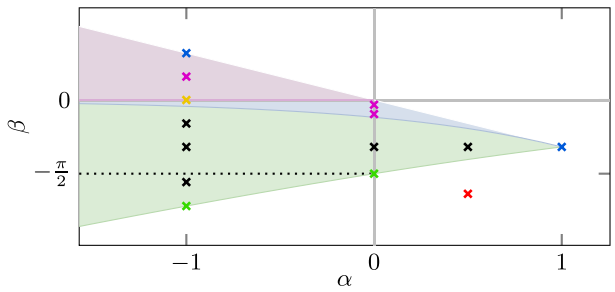

FIG. 4. Eigenvalue phase diagram for Eq. (33), reproduced from Ref. 26, where $\alpha = a\tau$ and $\beta = b\tau$. Colored regions indicate a system for which a stationary solution exists, with purple, blue, and green corresponding to systems with, respectively, 1, 2, and 0 real eigenvalues; as shown in Ref. 26, these regions are unaffected by q. The $\times$ marks relate to the rows of Table I; their color indicates whether the corresponding system is unstable (red), is on a stability boundary (blue, green), has real eigenvalues (purple) or has no delay (orange).

C. Towards a treatment of initial conditions

One of the frequently ignored challenges in the study of DDEs is the treatment of initial conditions. A common approach is to choose a particular initial function (e.g., a constant function over $[0, \tau_{max}]$) and otherwise ignore the effect of changing it on the dynamics.

The dimensional hierarchy presented in Sec. V B provides a natural way of addressing this issue. By defining an initial function in terms of the amplitudes $c_n(0)$ along the most dominant modes, we have a finite set of numbers that can be varied to scan in a meaningful way the full breadth of initial conditions. Since each mode corresponds to a characteristic frequency, the number of modes will determine the temporal resolution with which we can analyze the solution.

Moreover, the initial conditions may not even be deterministic. For example, when dealing with real-world data, it is often the case that we do not know exactly what the initial conditions of a process should be, but have some idea of the range within which they fall. This can be accounted for by simply replacing the $c_n(0)$ by random variables $C_n(0)$.

We present now an example of how one might treat a probabilistic initial function, by characterizing the effect of noise on the collapse of Eq. (33) to its dominant mode(s) when $\alpha = \beta = \tau = 1$. To avoid limiting our result to a particular initial function, we set the probability distribution on the first four mode pairs of the process such that their power follows a normal distribution which is truncated at zero

$$|C_{|n|}(0)|^2 \overset{*}{\sim} \mathcal{N}(\mu_{|n|}, s^2_{|n|}), \quad |n| = 1, \ldots, 4, \tag{34}$$

and set the amplitudes of the other modes to zero. (We use $\overset{*}{\sim}$ to indicate that the distribution is only approximately normal; chosen values of μ_n, s_n are listed in Appendix A.) We then computed the probability distribution of the time $t_{0.95}$ for which the relative power in Eq. (31) is $RP(1, t_{0.95})$

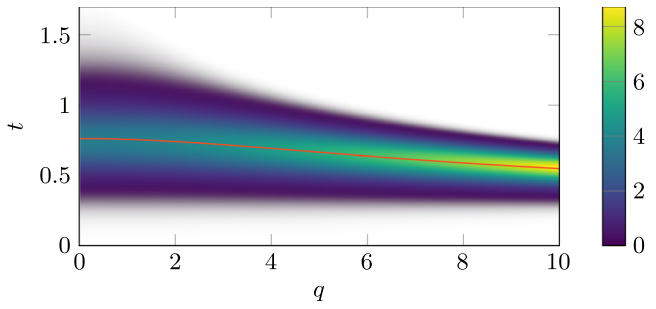

FIG. 5. Time required for 95% of power to collapse onto the first eigenmode pair. The system under study is given by Eq. (2), with $\kappa(\tau) = \delta(\tau) + \delta(\tau - 1)$. The initial condition is a probability density over the four most dominant modes, given by Eq. (34) and Appendix A. Color intensity indicates the probability that the system, for a given noise strength q, will have 95% of its power concentrated in mode C_0 at time t; relative power is as defined in Eq. (31). The orange line indicates the mean of the distribution. We see that by increasing the amount of noise, we can guarantee collapse to C_0 in a shorter amount of time, which amounts to reducing the dimensionality of the dynamics. To obtain this result, we generated 400 000 different initial conditions, solved $RP(0, t) = 0.95$ numerically for each initial condition and noise level and binned the results into histograms.

$= 0.95$ (i.e., the time after which the first two conjugate modes capture 95% of the power). The result, shown in Fig. 5, shows that noise mostly tightens the distribution of $t_{0.95}$ around its mean.

VI. CONCLUSION

The central result of this work is the solution of the linear stochastic delay system of Eq. (2), in the form of the time-dependent p.d.f. given by Eqs. (18); we obtained this by showing that the chain method has a well-defined limit when the number of equations goes to infinity. This provides a practical expression which obviates the need to resort to Monte Carlo evaluations for linearizable SDDEs with distributed delay. It also demonstrates a viable approach to addressing the dimensionality problem of SDDEs: By passing to an intermediate finite-size matrix representation, we can exploit known results for multivariate SDEs, while still being able to write the final result in continuous time. Moreover, the discretized system is often more computationally tractable. It can be readily diagonalized to obtain good approximations of the eigenvalues, eigenfunctions, and associated dual functionals.

Since we require only that the discretized operator $\boldsymbol{A}^{(N)}$ be full-rank, our results are applicable to a larger parameter space than other current methods: No assumption on the magnitude of the delayed feedback or the presence of a bifurcation is imposed. In addition, the size of the low-dimensional representation we obtain is adaptable, and we provide an analytical expression to estimate its error.

Although our treatment here has been limited to the linear SDDE with additive noise, Eq. (13b) should provide a good starting point for further extensions. For instance, multiplicative noise could be included by appropriately modifying $\boldsymbol{Q}^{(N)}$ and solving the resulting fully linear SDE. Furthermore, multiple approaches have already been proposed for extending the chain method to nonlinear DDEs,[17,22] and so, it is reasonable to suspect that the same could be done in this case for SDDEs.

Of particular interest is the fact that the series solution we obtained in Eqs. (18) can in some cases be well approximated with only a few terms. This not only makes it convenient for computational applications but also suggests that solutions are in fact effectively low-dimensional. Indeed, similar to how nonlinear DDEs may collapse to an attractor with fractal dimension,[10,11] we found that solutions to SDDEs collapse to a few dominant modes, leaving the other modes with vanishingly small relative power. In contrast to centre-manifold approaches, this low-dimensional representation is not limited to parameter regimes around a bifurcation. It may thus allow for a tractable definition of probability densities for SDDEs and thereby permit a more general treatment of the role played by initial conditions.

ACKNOWLEDGMENTS

The authors thank the Fonds de Recherche du Québec—Nature et Technologies and the Natural Sciences and Engineering Research Council of Canada for financially supporting this research.

APPENDIX A: NUMERICAL PARAMETERS FOR THE PROBABILISTIC INITIAL CONDITION

The parameters of Eq. (34) were set with

$$\mu_n = -7.5\frac{q^2|\mathcal{N}_n|^2}{\mathrm{Re}(\lambda_n)}\left[\tanh\left(\frac{3.4-n}{0.4}\right)+1\right],$$
$$\sigma_n = \sqrt{\mu_n/3},$$

which for the case $\alpha = \beta = \tau = 1$ works out to

$$\mu_1 = 4.820, \quad \mu_2 = 4.816, \quad \mu_3 = 4.246, \quad \mu_4 = 0.229,$$
$$\sigma_1 = 1.268, \quad \sigma_2 = 1.267, \quad \sigma_3 = 1.190, \quad \sigma_4 = 0.276\,.$$

APPENDIX B: OBTAINING THE EIGENFUNCTION DECOMPOSITION THROUGH A DISCRETIZED REPRESENTATION

In Sec. II, we stated some classic results for linear FDEs. The goal of this appendix is to recover these expressions for the dynamical system Eq. (2) of interest here, in the case where $q = 0$. Specifically, we will derive the expressions for the eigenvalues and eigenvectors by means of an intermediate discretization step. As seen in Sec. III, the discretized system turns out to be a good starting point for extending the results to stochastic equations.

Recall that we defined x to be a function over a finite domain $[0, T]$ evolving in time such that (c.f. Eq. (4))

$$x_\tau(t) \stackrel{\text{def}}{=} x_0(t-\tau) \quad \tau \in [0,T]. \tag{B1}$$

For legibility, this Appendix sets $T \equiv \tau_{max}$. Following the chain method,[4,31] we approximate the function $x(t)$ with a vector by sampling it at regular intervals

$$x_k^{(N)}(t) \equiv x_{\tau_k}(t), \tag{B2}$$

where

$$\tau_k = \frac{kT}{N}, \quad k \in \mathbb{N}, \quad 0 \le k < N\,. \tag{B3}$$

The generator $\boldsymbol{A}$ of Eq. (5) then becomes the matrix $\boldsymbol{A}^{(N)}$ defined in Eq. (13b)

$$\boldsymbol{A}^{(N)} \stackrel{\text{def}}{=} \begin{pmatrix} \bar\kappa(0)\frac{T}{N} & \bar\kappa\left(\frac{T}{N}\right)\frac{T}{N} & & \bar\kappa\left(\frac{(N-1)}{N}T\right)\frac{T}{N} \\ N/T & -N/T & & \\ 0 & N/T & -N/T & \\ & & \ddots & \\ & & N/T & -N/T \end{pmatrix}. \tag{B4}$$

Eigenvalues and vectors are obtained either by solving

$$(\lambda_n I - \boldsymbol{A}^{(N)})\boldsymbol{\phi}_n^{(N)} = 0, \tag{B5a}$$

for the right eigenvectors, or

$$(\lambda_n I - \boldsymbol{A}^{(N)})^{\mathsf{T}} (\phi_n^{(N)\dagger})^{\mathsf{T}} = 0 \tag{B5b}$$

for the left eigenvectors. Below, we derive the solution to Eq. (B5a); the method is exactly analogous for Eq. (B5b).

Using elementary row operations, and assuming[29] that $\lambda_n \neq -N/T$, we put $\lambda_n I - \boldsymbol{A}^{(N)}$ in echelon form by dividing each line except the first by N/T:

$$\lambda_n I - \boldsymbol{A}^{(N)} \sim \begin{pmatrix} \lambda_n - \kappa(0)\frac{T}{N} & -\kappa\left(\frac{T}{N}\right)\frac{T}{N} & -\kappa\left(\frac{2T}{N}\right)\frac{T}{N} & \cdots & -\kappa(\frac{(N-1)T}{N})\frac{T}{N} \\ -1 & \frac{\lambda_n T}{N}+1 & 0 & & 0 \\ 0 & -1 & \frac{\lambda_n T}{N}+1 & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & 0 & \cdots & -1 & \frac{\lambda_n T}{N}+1 \end{pmatrix}.$$

Multiplying the second row by $(\frac{\lambda_n T}{N}+1)^{-1}$ and adding to the third, we get

$$\sim \begin{pmatrix} \lambda_n - \kappa(0)\frac{T}{N} & -\kappa\left(\frac{T}{N}\right)\frac{T}{N} & -\kappa\left(\frac{2T}{N}\right)\frac{T}{N} & \cdots & -\kappa(\frac{(N-1)T}{N})\frac{T}{N} \\ -1 & \frac{\lambda_n T}{N}+1 & 0 & & 0 \\ -\left(\frac{\lambda_n T}{N}+1\right)^{-1} & 0 & \frac{\lambda_n T}{N}+1 & & \vdots \\ \vdots & & & \ddots & 0 \\ 0 & 0 & \cdots & -1 & \frac{\lambda_n T}{N}+1 \end{pmatrix}.$$

Repeating this process all the way down, and then multiplying each row by $(\frac{\lambda_n T}{N}+1)^{-1}$, finally yields

$$\sim \begin{pmatrix} \lambda_n - \sum_{k=0}^{N-1} \kappa\left(\frac{kT}{N}\right)\left(\frac{\lambda_n T}{N}+1\right)^{-k}\frac{T}{N} & 0 & 0 & \cdots & 0 \\ -\left(\frac{\lambda_n T}{N}+1\right)^{-1} & 1 & 0 & & 0 \\ -\left(\frac{\lambda_n T}{N}+1\right)^{-2} & 0 & 1 & & \vdots \\ \vdots & & & \ddots & 0 \\ -\left(\frac{\lambda_n T}{N}+1\right)^{-(N-1)} & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{B6}$$

To obtain non-trivial solutions to Eq. (B5a), the rows of the matrix in Eq. (B6) must be linearly dependent. This in turn implies

$$\lambda_n = \sum_{k=0}^{N-1} \kappa\left(\frac{kT}{N}\right)\left(1+\frac{\lambda_n T}{N}\right)^{-k}\left(\frac{T}{N}\right), \tag{B7a}$$

which is $\boldsymbol{A}^{(N)}$'s characteristic equation. Substituting Eqs. (B7a) back into Eq. (B6), the components of the right eigenvectors are seen to be

$$\phi_{n,i}^{(N)} = \left(1+\frac{\lambda_n T}{N}\right)^{-i}, \quad (0 \leq i < N). \tag{B7b}$$

Repeating the same diagonalization procedure for Eq. (B5b), we find the components of the left eigenvectors

$$\phi_{n,i}^{(N)\dagger} = \begin{cases} \mathcal{N}_n^{(N)}, & i = 0 \\ \mathcal{N}_n^{(N)} \sum_{k=i}^{N-1} \kappa\left(\frac{kT}{N}\right)\left(1+\frac{\lambda_n T}{N}\right)^{i-k-1}\left(\frac{T}{N}\right)^2, & i \geq 1, \end{cases} \tag{B7c}$$

where the normalization constant $\mathcal{N}_n^{(N)}$ is given by

$$\mathcal{N}_n^{(N)} = \left[1+\sum_{k=1}^{N-1} k\cdot\kappa\left(\frac{kT}{N}\right)\left(1+\frac{\lambda_n T}{N}\right)^{-k-1}\left(\frac{T}{N}\right)^2\right]^{-1}. \tag{B7d}$$

The constant $\mathcal{N}_n^{(N)}$ is chosen such that

$$\boldsymbol{\phi}_n^{(N)\dagger}(\boldsymbol{\phi}_m^{(N)}) = \delta_{n,m} \equiv \begin{cases} 1, & \text{if } n = m, \\ 0, & \text{if } n \neq m. \end{cases} \tag{B7e}$$

To revert to the continuous representation and recover the expressions in Sec. II, we use Eq. (B3) to replace either k or i with $\frac{\tau N}{T}$. Then

$$\phi_{n,\tau}^{(N)} = \lim_{N\to\infty}\left(1+\frac{\lambda_n T}{N}\right)^{-\tau N/T} = e^{-\lambda_n \tau}, \tag{B8}$$

which is consistent with Eq. (6). Recognizing Eqs. (B7a), (B7c), and (B7d) as Riemann sums, we also find

$$\lambda_n = \int_0^T \kappa(\tau) e^{-\lambda_n \tau} d\tau, \tag{B9}$$

$$\phi_n^\dagger x = \mathcal{N}_n x_0 + \mathcal{N}_n \int_{0^+}^T d\tau \int_t^T d\nu\, \kappa(\nu) e^{\lambda_n(\tau-\nu)} x_\tau, \tag{B10}$$

$$\mathcal{N}_n = \left[1+\int_{0^+}^T \kappa(\tau)\tau e^{-\lambda_n \tau}\, dt\right]^{-1}. \tag{B11}$$

APPENDIX C: CHOOSING A TRUNCATION SCHEME

To evaluate Eq. (18d) in practice, one needs to select a truncation scheme. For example, we might choose to evaluate $\Sigma_{\tau\nu}(t)$ as

$$\Sigma_{\tau\nu}(t) \approx \sum_{n=1}^{n_{max}}\sum_{m=1}^{n_{max}}\Big(\Phi_{\tau,\nu,t}(n,m) + \Phi_{\tau,\nu,t}(-n,m) + \Phi_{\tau,\nu,t}(n,-m) + \Phi_{\tau,\nu,t}(-n,-m)\Big) \tag{C1a}$$

where n_{max} is some suitably large number. However, we could just as well instead choose to compute

$$\Sigma_{\tau\nu}(t) \approx \sum_{k=2}^{k_{max}}\sum_{n=1}^{k-1}\Big(\Phi_{\tau,\nu,t}(n,k-n) + \Phi_{\tau,\nu,t}(-n,k-n) + \Phi_{\tau,\nu,t}(n,-k+n) + \Phi_{\tau,\nu,t}(-n,-k+n)\Big) \tag{C1b}$$

with a suitable k_{max}. Both these schemes converge to the true autocovariance in the limit $n_{max}, k_{max} \to \infty$, but their behavior for finite n_{max} and k_{max} can differ substantially.

To illustrate this, we consider the following equation:

$$\Sigma_{\tau\tau}(\infty) = q^2 \sum_{n,m} \mathcal{N}_n \mathcal{N}_m \frac{-e^{-(\lambda_n+\lambda_m)\tau}}{\lambda_n+\lambda_m}. \tag{C2}$$

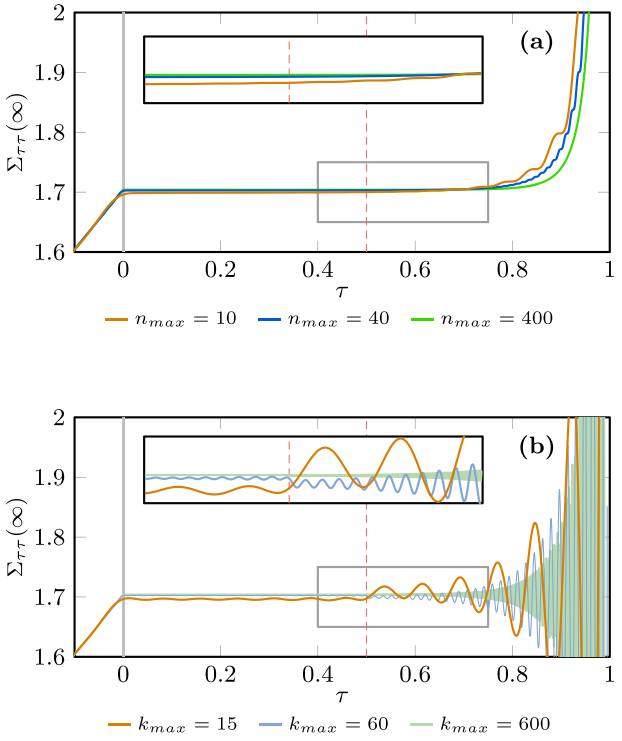

FIG. 6. Evaluation of the test series (C2) at various time lags, using either (a) Eq. (C1a) or (b) Eq. (C1b). As shown in Appendix E 4, the series converges absolutely for τ between 0 and 0.5; the latter bound is indicated by the dashed red line. The values of n_{max} and k_{max} were chosen to ensure the curves of the same colour are computed using a comparable number of terms, with the curves in (b) including a few more. The scheme given by Eq. (C1a) is seen to converge much more quickly and reliably, even when the series is only conditionally convergent. The value of the series veers away from $\mathrm{Var}(X_0(\infty))$ when $\tau < 0$ as it is no longer interpretable as a variance, a behavior which contrasts with Fig. 1, where ν is fixed to 0. These series were computed for a kernel $\kappa(\tau) = -\delta(\tau-1)$ and noise amplitude $q = 1$; initial condition is inconsequential, as we are considering the steady state.

This is almost identical to Eq. (25), with the difference that we set $\nu = \tau$ rather than $\nu = 0$. Since $\Sigma_{\tau\tau}(\infty)$ computes the stationary state variance, it should in fact be independent of τ. Thus, the flatness of $\Sigma_{\tau\tau}(\infty)$ with respect to t provides a measure of the convergence quality.

Figure 6 shows the evaluation of Eq. (C2) using Eq. (C1a) [Fig. 6(a)] and Eq. (C1b) [Fig. 6(b)]; evidently, the former provides much better convergence behavior in this case. The reason for this is that the terms of the covariance series (18d) form a periodic pattern on the (n, m) grid,[30] and Eq. (C1a) does a better job of selecting terms which cancel each other's deviation from the limit. Because the pattern depends on the chosen value of τ and ν, so too will the "optimal" summation scheme; nevertheless, we found Eq. (C1a) to be sufficiently well-behaved for the cases presented in this paper. For a more in-depth discussion on the differences between summation schemes, see Sec. 4.4 of Ref. 30.

APPENDIX D: DERIVATION OF MODE POWER

Since Eq. (27) is Markovian, we can use a standard technique to obtain from it the mean amplitude:

$$\begin{aligned} d\langle C_n\rangle &= \langle dC_n\rangle = \langle \lambda_n C_n\, dt + q\mathcal{N}_n\, dW\rangle = \lambda_n\langle C_n\rangle dt \\ \langle C_n(t)\rangle &= e^{\lambda_n t}\langle C_n(0)\rangle\,. \end{aligned} \tag{D1}$$

Similarly, the mean power is obtained as

$$\begin{aligned} |C_n(t+dt)|^2 &= C_n^*(t+dt)C_n(t+dt) \\ &= \left[C_n^*(t) + \lambda_n^* C_n^*(t)dt + q\mathcal{N}_n^* dW\right] \\ &\quad \times \left[C_n(t) + \lambda_n C_n(t)dt + q\mathcal{N}_n dW\right] \\ &= C_n^*(t)C_n(t) + (\lambda_n + \lambda_n^*)C_n^*(t)C_n(t)dt \\ &\quad + q^2\mathcal{N}_n^*\mathcal{N}_n dW^2 + (C_n^*\mathcal{N}_n + C_n\mathcal{N}_n^*)dW \\ &\quad + \mathcal{O}(dt\, dW) \end{aligned}$$

$$\begin{aligned} d|C_n(t)|^2 &\overset{\text{def}}{=} C_n^*(t+dt)C_n(t+dt) - C_n^*(t)C_n(t) \\ \frac{d}{dt}\langle|C_n(t)|^2\rangle &= 2\mathrm{Re}(\lambda_n)\langle|C_n(t)|^2\rangle + q^2\,|\mathcal{N}_n|^2 \\ \langle|C_n(t)|^2\rangle &= e^{2\mathrm{Re}(\lambda_n)t}\left(\langle|C_n(0)|^2\rangle + \frac{q^2\,|\mathcal{N}_n|^2}{2\mathrm{Re}(\lambda_n)}\right) - \frac{q^2\,|\mathcal{N}_n|^2}{2\mathrm{Re}(\lambda_n)}, \end{aligned} \tag{D2}$$

where we have used

$$\langle dW^2\rangle = dt \quad \text{and} \quad \langle dt^2\rangle = 0\,.$$

Here, * indicates complex conjugation.

APPENDIX E: CONVERGENCE DOMAIN OF THE AUTOCOVARIANCE

The goal in this Appendix is to justify Eq. (24), from which we deduced convergence conditions in the main text. We first introduce an asymptotic expression for the eigenvalues in Appendixes E 1 through E 3. This is then used in Appendix E 4 to sketch the derivation of Eq. (24). For full details, we refer the reader to AR's thesis,[30] specifically Secs. II B, IV D, and Appendix B.

We make the same assumptions as Sec. III regarding the spectrum of $\boldsymbol{A}$, namely, that only a finite number of its eigenvalues are real and that its eigenfunctions span C. The latter implies, in particular, that eigenvalues may have arbitrarily large negative real part.

1. Ordering of eigenvalues

The expressions below require us to choose an ordering for the eigenvalues λ_n, which we set as follows:

$$\begin{aligned} &|n| < |m| && \text{if } \lambda_n \in \mathbb{R} \text{ and } \lambda_m \in \mathbb{C}; \\ &|n| < |m| && \text{if } \mathrm{Re}(\lambda_n) > \mathrm{Re}(\lambda_m); \\ &\lambda_n \in \mathbb{R} \text{ iff } \lambda_{-n} \in \mathbb{R}\,; && \\ &\lambda_n > \lambda_{-n} > \lambda_{n+1} && \text{if } \lambda_n, \lambda_{n+1} \in \mathbb{R}; \\ &\lambda_{-n} = \lambda_n^*, \mathrm{Im}(\lambda_n) > 0 && \text{if } \lambda_n \in \mathbb{C}. \end{aligned}$$

This choice of ordering highlights the fact that complex solutions to Eq. (7) come in conjugate pairs when the kernel κ is real-valued, by associating with each pair a single magnitude $|n|$. Crucially, this means that if the real solutions are even numbered, *then we leave* λ_0 *undefined*.

2. Single delay approximation

Under the above-stated assumptions, we can limit the asymptotic analysis of Eq. (18d) to SDDEs with a single delay at $\tau = \tau_{max}$. This can be justified by considering a Dirac comb kernel

$$\kappa_{\text{comb}}(\tau) = \sum_{i=0}^{M} b_i \delta(\tau - \tau_i), \tag{E1}$$

where M is finite but arbitrarily large. Such a comb can be used to approximate any continuous kernel of compact support—in fact, this is the approximation we use to construct Eq. (13b). With κ_{comb}, the characteristic Eq. (7) becomes

$$\sum_{i=0}^{M} b_i\, e^{-\lambda_n \tau_i} = \lambda_n\,. \tag{E2}$$

Now, let $\epsilon > 0$ be given, set $\tau_M = \tau_{max}$ and choose $i \in \mathbb{N}$, where $0 \le i < M$. Since the real parts of λ_n are assumed to take arbitrarily large negative values, and since $\tau_M - \tau_i$ is finite, there exists an $n^* \in \mathbb{N}$ such that

$$|b_i e^{-\lambda_n \tau_i}| \le \frac{\epsilon}{M}|b_M e^{-\lambda_n \tau_M}|, \quad \forall n \ge n^*.$$

Repeating this process for each $i < M$, we get

$$\sum_{i=0}^{M-1} |b_i\, e^{-\lambda_n \tau_i}| \le \epsilon |b_M\, e^{-\lambda_n \tau_M}|, \quad \forall n \ge n^*,$$

which, when substituting into Eq. (E2), leads to

$$b_M\, e^{-\lambda_n \tau_{max}} \le \lambda_n \le (1+\epsilon) b_M\, e^{-\lambda_n \tau_{max}}, \quad \forall n \ge n^*\,.$$

And since ϵ can be chosen arbitrarily small, we can state that

$$b_M\, e^{-\lambda_n \tau_{max}} \approx \lambda_n, \tag{E3}$$

for n large enough.

Thus, we see that asymptotically, the characteristic equation—and therefore the distribution of eigenvalues—approaches that of a single delay DDE.

3. Asymptotic eigenvalue distribution

A single delay SDDE takes the form

$$\frac{dx_0(t)}{dt} = b x_0(t - \tau_{max}) + q\, dW\,. \tag{E4}$$

The eigenvalue spectrum of this equation is independent of q, so we can refer to results derived for DDEs. In this case, it is known that the eigenvalues are given by the Lambert W function:[9,12]

$$\lambda_n = \frac{W_n(b\tau_{max})}{\tau_{max}}\,. \tag{E5}$$

For large n, the Lambert function admits the following asymptotic expression:

$$W_n(\beta) \approx -\log\frac{2\pi(|n|-1)}{|\beta|} + i\left(2\pi(n-\mathrm{sgn}(n)) - \mathrm{sgn}(\beta)\frac{\pi}{2}\right), \tag{E6}$$

which can be obtained from Ref. 35 [Eqs. (4.4) and (4.6)] by restricting β to the reals. A similar expression is also found in Ref. 9 [Eq. (4.20)]. A more direct derivation making use of the fact that β is real is given in Ref. 30 (Sec. 2.2).

Combining Eqs. (E6) and (E5), we find an asymptotic expression for the eigenvalues of Eq. (E4):

$$\lambda_n \sim -\frac{1}{\tau_{max}}\log\left|\frac{2\pi(n-1)}{b\,\tau_{max}}\right| + \frac{i}{\tau_{max}}\left(2\pi(n-1) - \mathrm{sgn}(b)\frac{\pi}{2}\right). \tag{E7}$$

4. Asymptotic analysis of the covariance series

We can now analyze the absolute convergence of the autocovariance series in Eq. (18d). We check for *absolute* convergence for two reasons: first because it is analytically more tractable than simple convergence, and second because it ensures that we need not worry about specifying the order in which the double sum is taken.

We recall the definition of $\Phi_{\tau,\nu,t}(n,m)$ introduced in Eq. (24),

$$\Phi_{\tau,\nu,t}(n,m) = q^2\mathcal{N}_n\mathcal{N}_m\frac{e^{(\lambda_n+\lambda_m)t}-1}{\lambda_n+\lambda_m}e^{-\lambda_n t}e^{-\lambda_m\nu},$$

which allows us to rewrite Eq. (18d) more compactly as

$$\Sigma_{\tau\nu}(t) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\Phi_{\tau,\nu,t}(n,m)\,. \tag{E8}$$

Our objective then is to show that

$$\sum_{n,m}|\Phi_{\tau,\nu,t}(n,m)| \sim \sum_{|n|,|m|<n'}|\Phi_{\tau,\nu,t}(n,m)| + B_{\tau,\nu}\sum_{k=n'}^{\infty}\frac{\log k}{k^{2-(\tau+\nu)/\tau_{max}}}, \tag{E9}$$

for some n' large enough.

a. Properties of the covariance series terms

If n'' is the lowest positive index for which $\lambda_{n''}$ is complex, then given the ordering defined in Appendix E 1, we have the following properties for the terms $\Phi_{\tau,\nu,t}(n,m)$:

$$\begin{aligned}
&\Phi_{\tau,\nu,t}(n,m) = \Phi_{\tau,\nu,t}(m,n), && \forall\, n,m\\
&\Phi_{\tau,\nu,t}(n,-m) = \Phi_{\tau,\nu,t}(n,m)^*, && \forall\, |n|<n'', |m|\geq n''\\
&\Phi_{\tau,\nu,t}(n,m) = \Phi_{\tau,\nu,t}(-n,-m)^*, && \forall\, |n|,|m|\geq n''\,.
\end{aligned} \tag{E10}$$

Given these properties, one can show that the imaginary parts of Eq. (E8) cancel out, as they should. In what follows, we impose that $0\leq n''\leq n'<\infty$, where the last inequality is permitted due to the assumption that there are only finitely many real eigenvalues.

b. Separation of the series terms

If we take the absolute value of each term in Eq. (E8) and employ Eq. (E10), the series can be ordered as

$$\sum_{n,m}|\Phi_{\tau,\nu,t}(n,m)| = S_1+S_2+S_3+S_4, \tag{E11}$$

where

$$\begin{aligned}
S_1 &= \sum_{|n|,|m|<n'}|\Phi_{\tau,\nu,t}(n,m)|,\\
S_2 &= 2\sum_{|n|<n'}\sum_{|m|\geq n'}|\Phi_{\tau,\nu,t}(n,m)|,\\
S_3 &= \sum_{n\geq n',m\geq n'}|\Phi_{\tau,\nu,t}(n,m)| + \sum_{n\geq n',m\geq n'}|\Phi_{\tau,\nu,t}(-n,-m)|,\\
S_4 &= \sum_{n\geq n',m\geq n'}|\Phi_{\tau,\nu,t}(n,-m)| + \sum_{n\geq n',m\geq n'}|\Phi_{\tau,\nu,t}(-n,m)|\,.
\end{aligned}$$

The series S_1 here is recognized as the first term of Eq. (E9) and requires no further analysis. The series S_3 (S_4) collects the terms where n and m are both large and of the same (opposite) sign.

Asymptotically, the terms of S_2, S_3, and S_4 behave as

$$|\Phi_{\tau,\nu,t}(n,m)| \sim \frac{q^2|e^{(\lambda_n+\lambda_m)t}-1||e^{-\lambda_n\tau}||e^{-\lambda_m\nu}|}{|\lambda_n+\lambda_m||1+\lambda_n\tau_{max}||1+\lambda_m\tau_{max}|}\,. \tag{E12}$$

Substituting Eq. (E7) into Eq. (E12), we can show that

$$S_2 \lesssim 4n'B^{\mathrm{I}}_{\tau,\nu,n'}\sum_{m\geq n'}\frac{1}{m^{2-\nu/\tau_{max}}}, \tag{E13}$$

$$S_3 \lesssim 2B^{\mathrm{III}}_{\tau,\nu}\sum_{k=n'}^{\infty}\frac{\log k}{k^{2-(\tau+\nu)/\tau_{max}}}, \tag{E14}$$

$$S_4 \sim 4B^{\mathrm{IV}}_{\tau,\nu}\sum_{k=n'}^{\infty}\frac{\log k}{k^{2-(\tau+\nu)/\tau_{max}}}, \tag{E15}$$

where B^{I}, B^{III}, and B^{IV} are constants. Summing these while keeping only the leading order terms gives us Eq. (E9). For illustration, we outline below the procedure by which we obtained Eq. (E14); the procedure for Eq. (E15) is more technical but broadly similar. Full details for all three series can be found in Appendix B3 of Ref. 30.

c. Derivation of Eq. (E14)

To treat the double sum in S_3, we need to set a summation order. We choose to collect the terms for which $n+m=k$ for some k, following the scheme given by Eq. (C1b). Since S_1 and S_2 already cover all terms where either n or m is smaller than n', the smallest possible value for k is $n'+n'=2n'$. Consequently the S_3 series can be written

$$\sum_{k=2n'}^{\infty}\sum_{n=n'}^{k-n'}|\Phi_{\tau,\nu,t}(n,k-n)| + \sum_{k=2n'}^{\infty}\sum_{n=n'}^{k-n'}|\Phi_{\tau,\nu,t}(-n,-k+n)|.$$

We can find an asymptotic expression for the entire inner sum and treat that as a single term in the series to determine its convergence domain. Both double sums are analogous, so we only need to treat one; we will do the one on the left.

We want to determine the terms of Eq. (E12) for $n' \ll k$ and $n' \leq n \leq k - n'$. Since our choice for n' was arbitrary (it only needs to be finite), we make the additional assumption that $1 \ll n'$. This leads us to

$$|\lambda_n + \lambda_{k-n}| \sim \left| -\frac{1}{\tau_{max}} \log\left[\left(\frac{2\pi\, n}{b\, \tau_{max}}\right)\left(\frac{2\pi\,(k-n)}{b\,\tau_{max}}\right)\right] + \frac{i}{\tau_{max}}\,(2\pi\, n + 2\pi\,(k-n))\right| \sim \frac{2\pi\, k}{\tau_{max}} \tag{E16a}$$

$$|1 + \lambda_n \tau_{max}| \sim 2\pi n \tag{E16b}$$

$$|1 + \lambda_{k-n}\tau_{max}| \sim 2\pi(k-n) \tag{E16c}$$

$$|e^{-\lambda_n \tau}| \sim \left[\frac{2\pi\, n}{b\,\tau_{max}}\right]^{\tau/\tau_{max}} \tag{E16d}$$

$$|e^{-\lambda_{k-n}\nu}| \sim \left[\frac{2\pi(k-n)}{b\,\tau_{max}}\right]^{\nu/\tau_{max}} \tag{E16e}$$

$$|e^{(\lambda_n+\lambda_{k-n})t} - 1| \approx 1. \tag{E16f}$$

Substituting these into Eq. (E12), we get

$$\sum_{n=n'}^{k-n'} |\Phi_{\tau,\nu,t}(n,m)| \sim \sum_{n=n'}^{k-n'} \frac{\tau_{max}}{(2\pi)^3}\left(\frac{2\pi}{b\tau_{max}}\right)^{\tau+\nu/\tau_{max}} k^{-1} \times n^{-1+\tau/\tau_{max}}\,(k-n)^{-1+\nu/\tau_{max}} \sim B^{\mathrm{II}}_{\tau,\nu} \sum_{n=n'}^{k-n'} n^{-1+\tau/\tau_{max}}\,(k-n)^{-1+\nu/\tau_{max}}\,\frac{1}{k}.$$

If we define $x \equiv n/k$, then we can write this as

$$\sim B^{\mathrm{II}}_{\tau,\nu} k^{-2+\frac{\tau+\nu}{\tau_{max}}} \sum_{n=n'}^{k-n'} x^{-1+\frac{\tau}{\tau_{max}}}\,(1-x)^{-1+\frac{\nu}{\tau_{max}}}\,\frac{1}{k},$$

where we recognize a Riemann sum. In the limit of large k, we can replace it by an integral, giving us

$$\sim B^{\mathrm{II}}_{\tau,\nu} k^{-2+\frac{\tau+\nu}{\tau_{max}}} \int_{\frac{n'}{k}}^{1-\frac{n'}{k}} x^{-1+\frac{\tau}{\tau_{max}}}\,(1-x)^{-1+\frac{\nu}{\tau_{max}}}\,dx \tag{E17}$$

$$\sim B^{\mathrm{II}}_{\tau,\nu} k^{-2+\frac{\tau+\nu}{\tau_{max}}} \int_{0}^{1} x^{-1+\frac{\tau}{\tau_{max}}}\,(1-x)^{-1+\frac{\nu}{\tau_{max}}}\,dx. \tag{E18}$$

Here, we recognize the integral as the Beta function $B\left(\frac{\tau}{\tau_{max}}, \frac{\nu}{\tau_{max}}\right)$, which is finite for all strictly positive τ and ν. The result is that for $\tau, \lambda > 0$, S_3 can be written

$$\sum_{n\geq n', m\geq n'} |\Phi_{\tau,\nu,t}(n,m)| \sim B^{\mathrm{II}}_{\tau,\nu} B\left(\frac{\tau}{\tau_{max}}, \frac{\nu}{\tau_{max}}\right) \sum_{k=n'}^{\infty} \frac{1}{k^{2-(\tau+\nu)/\tau_{max}}}.$$

The case where either τ or ν is zero must be solved somewhat differently: In this case, the series' terms are asymptotically proportional to $\frac{\log k}{k^{2-(\tau+\nu)/\tau_{max}}}$ (see Ref. 30, p. 83). Thus, to capture all cases we write

$$\sum_{n\geq n', m\geq n'} |\Phi_{\tau,\nu,t}(n,m)| \lesssim B^{\mathrm{III}}_{\tau,\nu} \sum_{k=n'}^{\infty} \frac{\log k}{k^{2-(\tau+\nu)/\tau_{max}}}, \tag{E19}$$

where all constants were absorbed into $B^{\mathrm{III}}_{\tau,\nu}$. Treating the other series in S_3 the same way yields the factor of 2 in Eq. (E14).

[1] A. Amann, E. Schöll, and W. Just, Phys. A **373**, 191 (2007).
[2] D. J. Amit and N. Brunel, Network: Comput. Neural Syst. **8**, 373 (1997).
[3] A. Bahar and X. Mao, J. Math. Anal. Appl. **292**, 364 (2004).
[4] H. T. Banks, *Nonlinear Systems Applications* (Elsevier, 1977), p. 21.
[5] R. Bellman, *Differential-Difference Equations*, Mathematics in Science and Engineering Vol. 6 (Academic Press, New York, 1963).
[6] D. Bratsun, D. Volfson, L. S. Tsimring, and J. Hasty, Proc. Natl. Acad. Sci. U.S.A. **102**, 14593 (2005).
[7] G. Burgers, Clim. Dyn. **15**, 521 (1999).
[8] P. Chang, L. Ji, H. Li, and M. Flügel, Phys. D **98**, 301 (1996).
[9] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, Adv. Comput. Math. **5**, 329 (1996).
[10] B. Dorizzi, B. Grammaticos, M. Le Berre, Y. Pomeau, E. Ressayre, and A. Tallet, Phys. Rev. A **35**, 328 (1987).
[11] J. Doyne Farmer, Phys. D **4**, 366 (1982).
[12] T. Erneux, *Applied Delay Differential Equations*, Surveys and Tutorials in the Applied Mathematical Sciences Vol. 3 (Springer, New York, 2009).
[13] T. D. Frank, Phys. Lett. A **380**, 1341 (2016).
[14] T. D. Frank and P. J. Beek, Phys. Rev. E **64**, 021917 (2001).
[15] J. García-Ojalvo and R. Roy, Phys. Lett. A **224**, 51 (1996).
[16] S. Guillouzic, I. L'Heureux, and A. Longtin, Phys. Rev. E **59**, 3970 (1999).
[17] I. Györi, Comput. Math. Appl. **16**, 195 (1988).
[18] J. K. Hale and S. M. V. Lunel, *Introduction to Functional Differential Equations*, 1st ed., Applied Mathematical Sciences Vol. 99 (Springer-Verlag, New York, 1993).
[19] A. Hutt and J. Lefebvre, Markov Processes Relat. Fields **22**(3), 555–572 (2016).
[20] T. Insperger and G. Stépán, IET Control Theory Appl. **1**, 553 (2007).
[21] L. Jaurigue, A. Pimenov, D. Rachinskii, E. Schöll, K. Lüdge, and A. G. Vladimirov, Phys. Rev. A **92**, 053807 (2015).
[22] B. Krasznai, I. Gyori, and M. Pituk, Math. Comput. Modell. **51**, 452 (2010).
[23] J. Lefebvre, A. Hutt, J.-F. Knebel, K. Whittingstall, and M. M. Murray, J. Neurosci. **35**, 2895 (2015).
[24] J. Lefebvre, A. Hutt, V. G. LeBlanc, and A. Longtin, Chaos: An Interdisciplinary J. Nonlinear Sci. **22**, 043121 (2012).
[25] J. Lei and M. Mackey, SIAM J. Appl. Math. **67**, 387 (2007).
[26] M. C. Mackey and I. G. Nechaeva, J. Dyn. Differ. Equations **6**, 395 (1994).
[27] J. Miękisz, J. Poleszczuk, M. Bodnar, and U. Foryś, Bull. Math. Biol. **73**, 2231 (2011).
[28] In general, C may be decomposed as $C = \overline{\mathcal{M} \oplus \mathcal{S}}$, where $\mathcal{M}$ is the space spanned by the generalized eigenfunctions and $\mathcal{S}$ is the space of "small solutions," which reach zero in finite time. In many cases $\mathcal{S} = 0$, and thus C is equal to the closure of $\mathcal{M}$. See Ref. 18 (Sec. 3.3).
[29] Since we want to be able to vary the coarseness of the discretization, any singularities arising only for particular values of N are irrelevant.
[30] A. René, "Spectral solution method for distributed delay stochastic differential equations," Master's thesis (Université d'Ottawa/University of Ottawa, 2016).
[31] I. M. Repin, J. Appl. Math. Mech. **29**, 254 (1965).
[32] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications*, 2nd ed., Springer Series in Synergetics Vol. 18 (Springer-Verlag, New York, 1996).
[33] H. Smith, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, 1st ed., Texts in Applied Mathematics Vol. 57 (Springer-Verlag New York, 2007).
[34] J. Touboul, J. Stat. Phys. **149**, 569 (2012); arXiv:1209.2596.
[35] E. M. Wright, Proc. R. Soc. Edinburgh, Sect. A: Math. Phys. Sci. **65**, 193 (1959).
[36] B. Yang and X. Wu, AIAA J. **36**, 2218 (1998).