Higher-order stochastic differential equations and the positive Wigner function

P. D. Drummond
Centre for Quantum and Optical Science, Swinburne University of Technology, Melbourne 3122, Australia
(Received 4 October 2017; published 4 December 2017)

General higher-order stochastic processes that correspond to any diffusion-type tensor of higher than second order are obtained. The relationship of multivariate higher-order stochastic differential equations with tensor decomposition theory and tensor rank is explained. Techniques for generating the requisite complex higher-order noise are proved to exist either using polar coordinates and γ distributions, or from products of Gaussian variates. This method is shown to allow the calculation of the dynamics of the Wigner function, after it is extended to a complex phase space. The results are illustrated physically through dynamical calculations of the positive Wigner distribution for three-mode parametric downconversion, widely used in quantum optics. The approach eliminates paradoxes arising from truncation of the higher derivative terms in Wigner function time evolution. Anomalous results of negative populations and vacuum scattering found in truncated Wigner quantum simulations in quantum optics and Bose-Einstein condensate dynamics are shown not to occur with this type of stochastic theory.

DOI: 10.1103/PhysRevA.96.062104

I. INTRODUCTION

The Wigner distribution [1] of quantum mechanics is generically nonpositive. Hence, it cannot always be treated with probabilistic sampling methods. These negative Wigner distributions are caused dynamically by higher-order derivatives occurring in their dynamical evolution equations [2]. As a result, this issue is fundamentally related to the computational complexity of quantum mechanics [3], Bell inequality violations, and the problem of quantum simulation. It is one reason why probabilistic sampling is nontrivial for quantum dynamics.

Positive extensions of the Wigner function are known to exist [4–7]. These use dimension-doubled phase-space methods similar to the normally ordered positive-P distribution in quantum optics [8,9]. This paper investigates the fundamental question: Can one simulate these positive Wigner functions probabilistically?

To answer this question, a theory of higher-order stochastic differential equations (HSDE) equivalent to arbitrary multivariate partial differential equations is derived. This makes use of fundamental results in tensor decomposition theory. It extends previous work [10–14] by allowing the inclusion of arbitrary cross-derivative terms and complex variables. Methods for generating such higher-order noise are obtained. These results provide a dynamical basis for the positive Wigner function. The work unifies previous theoretical analyses in the physics literature, including earlier results for one-dimensional equations with higher-order polar noise [10,15] and Gaussian product noise [11–13].

These results are a step towards developing stochastic methods for Wigner functions. This would enable first-principles Wigner quantum field simulations using probabilistic sampling. Wigner representation methods have been applied to physical problems [16,17] with the high-order derivative terms truncated. This can cause errors [18,19], because the resulting theory is of a hidden variable type, closely related to stochastic electrodynamics. Proposals exist for adding weighted random corrections [20,21], but these can have sign problems with large numbers of modes. Other alternative phase-space methods for comparison purposes include the positive-P [9] and Gaussian [22] representations.

As a proof of principle, the HSDE method is applied to the problem of nonlinear parametric downconversion. This is used experimentally in quantum optics to generate squeezed states [23] and Bell violations [24]. The higher-order stochastic terms eliminate paradoxical results obtained when the Wigner dynamics is truncated. The new higher-order terms cancel unphysical vacuum depletion and third-order correlation effects [19,25–28]. These are known errors caused by Wigner truncation. Correct dynamical evolution is obtained when higher-order noise is included, as is demonstrated here with numerical simulations of coupled nonlinear parametric systems.

The positive Wigner distribution [4–7] is part of a larger class of Gaussian phase-space representations [22,29,30]. All such techniques may require higher-order derivatives, depending on the Hamiltonian and dissipative terms in nonlinear master equations, used for representing nonlinear loss in Bose-Einstein-condensate (BEC) systems [31]. HSDEs are therefore applicable to a variety of nonlinear physical problems. While this paper studies the fundamental properties of these equations, optimal numerical algorithms are still an open question. Important issues include noise optimization and stochastic gauge techniques that regularize instabilities in the extended phase-space equations [32]. The ungauged results obtained here are exact for constant coefficients but can only give results for short times with nonlinear coefficients.

The paper is organized as follows. In Sec. II the multidimensional Kramers-Moyal equation is introduced, together with its stochastic equivalent. In Sec. III, the stochastic identities are proved, together with a preliminary explanation of numerical algorithms. In Sec. IV, techniques for generating higher-order noise are explained. In Sec. V the positive Wigner function is defined, together with numerical examples in physics. Section VI gives a summary of the results.

II. MULTIVARIATE HIGHER-ORDER STOCHASTIC EQUATIONS

The Wigner function of quantum mechanics [1,2] is an example of a physical distribution equation that satisfies an

equation of motion with derivatives of arbitrarily high order, known as a Kramers-Moyal equation (KME) [33]. These equations occur widely in quantum and atom optics. Similar equations are found in many areas of physics and in other disciplines. Here a general KME is treated with cross derivatives of arbitrary order. It will be shown that there are complex stochastic processes that are equivalent to the KME solution in d real or complex space dimensions. The advantage of stochastic methods is that they are scalable to large numbers of modes.

A. Higher-order stochastic equations

This section will summarize the main results of the paper, for the reader who wishes to see where the derivations are heading. A diagonal case of a KME was treated in an earlier article [14]. Here a general partial differential equation having arbitrary spatial derivatives in d dimensions is treated. These generalize the Fokker-Planck equation [34,35] to give an nth-order tensor coefficient $D^{(n)}_{\mathbf{j}}$ [2,33], so that

$$\frac{\partial f}{\partial t} = \sum_{n=1}^{n_{\max}} \frac{(-1)^n}{n!} \sum_{\mathbf{j}} \left[\prod_{\mu=1}^{n} \frac{\partial}{\partial x_{j_\mu}} \right] D^{(n)}_{\mathbf{j}}(\mathbf{x}) f \,. \tag{2.1}$$

The vector $\mathbf{x}$ is a d-dimensional real or complex coordinate, and $f = f(\mathbf{x},t)$ is the distribution function, which is in general nonpositive. The total order of each derivative term is n, and the sum over $\mathbf{j} = j_1, \ldots j_n$ is a sum over all possible n-dimensional vectors of indices. For complex coordinates, $\partial/\partial x$ and $\partial/\partial x^*$ are treated as being independent, using Wirtinger calculus [36].

The stochastic process that corresponds to this equation relies on an expansion of the tensor coefficients $D^{(n)}_{\mathbf{j}}$ into sums of products of unary forms. A tensor can *always* be expanded as a sum of outer products [37]:

$$\mathbf{D}^{(n)}(\mathbf{x}) = \sum_{k=1}^{R} \prod_{m=1}^{M_k} [\mathbf{b}^{(n)}(\boldsymbol{k},\mathbf{x})]^{\otimes n_m(k)}, \tag{2.2}$$

where R is the number of different terms in the nth-order expansion. Given an nth-order tensor, its rank r is the minimal value of R. Here $\boldsymbol{k} = (k,m)$, and each vector $\mathbf{b}(\boldsymbol{k},\mathbf{x})$ enters the expansion with an integer power $n_m(k) \geqslant 0$. Since $\mathbf{D}^{(n)}$ is an nth-order tensor, it follows that $\sum_m n_m(k) = n$ for each term in the expansion. Determining the tensor rank r for large dimension d is an NP hard problem [38]. However, the results obtained here do not require the rank. Nonminimal decompositions are given in the following section.

The corresponding stochastic process is a sum over *complex* random noises $dw^{(\boldsymbol{n})}_{\mathbf{k}}$ in the direction of each vector $\mathbf{b}$. In general, each noise is a vector and has a vector order $\boldsymbol{n} = (n_1, \ldots n_{M_k})$. The stochastic process exists on a complex extended space, $\mathbf{z} = \mathbf{x} + i\mathbf{y}$, or a complex space with double the dimension if it was originally complex, provided the distribution over $\mathbf{z}$ has analytic moments. It is a sum over n,k,m, with a different term for each n:

$$d\mathbf{z} = \sum_{n,\boldsymbol{k}} \mathbf{b}^{(n)}(\boldsymbol{k},\mathbf{z}) dw^{(\boldsymbol{n}_k)}_{\boldsymbol{k}}, \tag{2.3}$$

where $\mathbf{b}^{(n)}(\boldsymbol{k},\mathbf{z})$ is an analytic continuation of $\mathbf{b}^{(n)}(\boldsymbol{k},\mathbf{x})$.

The vector noises of different k and n indices are independent. For the kth index there are $M = M_k$ vector noise components $dw^{(\boldsymbol{n})}_m$ with indices $m = 1 \ldots M$, that are correlated. The kth vector order, $\boldsymbol{n}_k = \boldsymbol{n}(k) = [n_1(k), \ldots n_M(k)]$, identifies the type of nth-order noise, $dw^{(\boldsymbol{n}_k)}_m$. For integer powers $p \geqslant 0$, the only nonvanishing analytic correlations of nth-order noise are

$$\left\langle \prod_{m=1}^{M} \left[dw^{(\boldsymbol{n})}_m \right]^{pn_m} \right\rangle = \frac{dt^p}{(n!)^p p!} \prod_{m=1}^{M} (pn_m)! \,. \tag{2.4}$$

The most elementary noise of this type has $\boldsymbol{n} = (1, \ldots 1)$, which will be termed multicomponent first-order noise. Algorithms that generate any dimension and order of these noises are provided in Sec. IV.

B. Elementary examples

1. Drift

The most obvious case is for $n_{\max} = 1$, when $D^{(1)}_j(\mathbf{x}) = A_j(\mathbf{x})$ is called the drift vector. The decomposition has one term with $n = 1$, so $dw^{(1)} = dt$. This gives the standard equation for characteristic solutions of a first-order partial differential equation:

$$d\mathbf{z} = \boldsymbol{A} dt. \tag{2.5}$$

2. Diffusion

The case with $n_{\max} = 2$ corresponds to a normal Ito stochastic differential equation [39,40]. One writes such equations by taking a matrix square root. This factorizes the diffusion matrix as $D^{(2)}_{ij}(\mathbf{x}) = \sum_{k=1}^{R} B_{ik}(\mathbf{x}) B_{jk}(\mathbf{x})$. This gives the form of Eq (2.2) if one defines $M_k = 1$, $n = n_1 = 2$, and $b^{(2)}_i(k) \equiv B_{ik}(\mathbf{x})$. The noise is the usual Gaussian noise with $\langle dw_k dw_{k'} \rangle = dt \delta_{kk'}$, where $d\boldsymbol{w} \equiv d\boldsymbol{w}^{(2)}$, and hence,

$$d\mathbf{z} = \boldsymbol{A} dt + \boldsymbol{B} d\boldsymbol{w}. \tag{2.6}$$

If there is negative diffusion, one obtains a complex matrix $\boldsymbol{B}$ [41]. In such dimension-doubled cases the equation may need to be regularized using stochastic gauges to prevent power-law distribution tails [42].

C. Higher-order stochastic equations

To give an understanding of higher-order stochastic methods, the approach will be illustrated with two different tensor decompositions, where the coefficients are assumed homogeneous.

1. Diagonal case

The first case considered is a generalization of the usual matrix square root method used in stochastic differential equations. Any diffusion tensor can always be made symmetric and then expanded in diagonal form, as a sum of tensor powers of rank-1 vectors:

$$\mathbf{D}^{(n)} = \sum_{k=1}^{R} \mathbf{b}^{(n)}(k)^{\otimes n} \,. \tag{2.7}$$

Here, $R \geqslant \mathrm{rank}_s(\mathbf{D}^{(n)})$, which is the *symmetric rank* of $\mathbf{D}^{(n)}$, and may be greater than the full rank. This approach is not always minimal [43] but it is simpler than the general decomposition. In this case $M_k = 1$, as there is just one vector in each decomposition, and $\boldsymbol{n} = n$. Defining the

matrix $\mathbf{B}^{(n)} = [\mathbf{b}^{(n)}(1), \ldots \mathbf{b}^{(n)}(R)]$ and the nth-order noise vector $d\mathbf{w}^{(n)} = (dw_1^{(n)}, \ldots dw_R^{(n)})^T$, the following nth-order stochastic equation is obtained:

$$d\mathbf{z} = \sum_n \mathbf{B}^{(n)} \cdot d\mathbf{w}^{(n)} . \tag{2.8}$$

Here the sum over decompositions k is treated as a matrix product. The noise terms $dw_k^{(n)}$ are independent nth-order noise terms whose only nonvanishing correlations were obtained in earlier work [10–14]:

$$\left\langle \left[dw_k^{(n)}\right]^{pn} \right\rangle = \frac{dt^p (pn)!}{(n!)^p p!} . \tag{2.9}$$

For third-order diffusion, this generalizes the Ito calculus form of a stochastic equation [41] to give a third-order noise equation in a complex extension of the real phase-space:

$$d\mathbf{z} = \mathbf{A}dt + \mathbf{B} \cdot d\mathbf{w} + \mathbf{C} \cdot d\mathbf{v}, \tag{2.10}$$

Here, $\boldsymbol{C} \equiv \mathbf{B}^{(3)} = (\mathbf{b}^{(3)}(1), \mathbf{b}^{(3)}(2) \ldots)$, while $d\mathbf{v} = d\mathbf{w}^{(3)}$ is a diagonal third-order noise, and the diffusion tensor is expanded as

$$\mathbf{D}^{(3)} = \sum_{k=1}^{R} [\mathbf{b}^{(3)}(k)]^{\otimes 3}. \tag{2.11}$$

The fundamental property of diagonal third-order noise [41] is that

$$\langle dv_i dv_j dv_k \rangle = dt \delta_{ik} \delta_{ij}. \tag{2.12}$$

While this allows a formal definition, taking the limit of $dt \to 0$ is not as simple as it is for second-order noise. This issue will be treated in later sections, with examples.

2. *Off-diagonal case*

An alternative expansion always exists even if a diagonal factorization is not known. Terms of the same coordinate derivative can be grouped together to give a diffusion tensor that is not symmetric. This is also generally not a minimal expansion, but it is simple to construct. For notational reasons, powers with $0 \leqslant n_m \leqslant n$ are allowed. Noises $dw_m^{(n)}$ with index $n_m = 0$ are set to zero, and it is assumed that $0^0 = 1$.

The *composite* order of each derivative term is $n = \sum_m n_m$, where $\mathbf{n} = (n_1, \ldots n_d)$ describes the *vectorial* order of each derivative. With this approach, each distinct vectorial order gives a different tensor decomposition. Each vectorial order then corresponds to a distinct term with different derivatives, each labeled $k = 1, \ldots R$ in the tensor decomposition of Eq. (2.2).

The vectors $\mathbf{b}^{(\boldsymbol{n})}(\boldsymbol{k}) = b_k^{(\boldsymbol{n})} \hat{\mathbf{x}}_m$ are proportional to unit vectors $\hat{\mathbf{x}}_m$ in the mth coordinate direction, and $b_k^{(\boldsymbol{n})} = (D_{\mathbf{j}})^{1/n}$, where $\mathbf{j}$ is the index of the kth nonvanishing tensor component. As an example, if the only term is $D_{111}^{(3)}$, the corresponding decomposition is such that the noise is of order $\boldsymbol{n} = 3$, and

$$\mathbf{D}^{(3)} = D_{111} \hat{\boldsymbol{x}}^{\otimes 3} . \tag{2.13}$$

This example shows the nonuniqueness of the decomposition. This example could also be treated as three first-order terms with a vector order $\mathbf{n} = (1,1,1)$, but the corresponding stochastic equation is likely to be sampled less efficiently.

As another example, if the only term is $D_{112}^{(3)}$, there are two correlated noises of vector order $\boldsymbol{n} = (2,1)$, with a decomposition into orthogonal vectors of

$$\mathbf{D}^{(3)} = D_{112} \hat{\boldsymbol{x}}^{\otimes 2} \hat{\boldsymbol{y}} . \tag{2.14}$$

The KME now has a form in which derivatives are grouped and summed over all R available derivative terms:

$$\frac{\partial f}{\partial t} = \sum_{n,k} \frac{1}{n!} \left[\prod_{m=1}^{d} \left(-\frac{\partial}{\partial x_m} \right)^{n_m} \right] \left(b_k^{(\boldsymbol{n})}\right)^n f . \tag{2.15}$$

In this case there is a corresponding higher-order stochastic process with an equation of motion given explicitly by

$$d\mathbf{z} = \sum_{n,\boldsymbol{k}} \mathbf{b}^{(n)}(\boldsymbol{k}) dw_{\boldsymbol{k}}^{(\boldsymbol{n}_k)}. \tag{2.16}$$

In this approach, the noise type $\boldsymbol{n}_k \equiv \boldsymbol{n}(k)$ depends on the term index k in the tensor expansion. It is notationally convenient to introduce one noise term per coordinate direction, with the convention that $dw_{k,m} = 0$ in any direction such that $n_m = 0$, which means that the noise-type vector $\boldsymbol{n}$ can be read off directly from the powers of the derivative terms in Eq. (2.15).

III. CHARACTERISTIC GREEN'S FUNCTION

To prove the higher-order stochastic differential equation (HSDE) results given in Eq. (2.3), characteristic functions for the KME and the corresponding stochastic process are equated. This results in a path-integral solution, following standard techniques first introduced by Wiener [44] and generalized by many other workers [45–48] in applications to diffusion processes.

A Green's function $g(\mathbf{x}, \Delta t | \mathbf{x}')$ in the original phase-space is defined to satisfy Eq. (2.1), with a δ-function initial condition of $\delta^d(\mathbf{x} - \mathbf{x}')$, assuming constant or slowly varying diffusion terms. Inserting the general decomposition expansion of Eq. (2.2),

$$\frac{\partial g}{\partial \Delta t} = \sum_{n,k,\mathbf{j}} \frac{(-1)^n}{n!} \left[\prod_{\mu=1}^{n} \frac{\partial}{\partial x_{j_\mu}} \right] \left\{ \prod_m [\mathbf{b}^{(n)}(\boldsymbol{k}, \boldsymbol{x})]^{\otimes n_m} \right\}_{\mathbf{j}} g . \tag{3.1}$$

The characteristic function for this is defined in the usual way as an expectation of a Fourier transform with respect to the difference coordinate, $\Delta x_j = x_j - x'_j$, so that if $d\mathbf{x}$ is a d-dimensional volume element,

$$\tilde{g}(\mathbf{q}, \Delta t | \mathbf{x}') = \int d\mathbf{x} g(\mathbf{x}, \Delta t | \mathbf{x}') e^{i\mathbf{q}\cdot\Delta\mathbf{x}} . \tag{3.2}$$

Two-dimensional example. As a simple example, consider a two-dimensional homogeneous case. A single term of this type might be

$$\frac{\partial g}{\partial \Delta t} = \frac{1}{2} \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} b_1 b_2 g. \tag{3.3}$$

After Fourier transforming, the resulting time-evolution equation for the characteristic Green's function in this simple case is therefore

$$\frac{\partial \tilde{g}}{\partial \Delta t} = \frac{1}{2} \int d^2\mathbf{x} e^{i\mathbf{q}\cdot\Delta\mathbf{x}} \frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} b_1 b_2 g . \tag{3.4}$$

Next, on partial integration, with the assumption of vanishing boundary terms,

$$\frac{\partial \tilde{g}}{\partial \Delta t} = \frac{1}{2}\int d^2\mathbf{x} b_1 b_2 g \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} e^{i\mathbf{q}\cdot\Delta\mathbf{x}}$$
$$= \frac{1}{2}\int d^2\mathbf{x} b_1 b_2 g(-q_1 q_2) e^{i\mathbf{q}\cdot\Delta\mathbf{x}}$$
$$= \frac{1}{2}(iq_1 b_1)(iq_2 b_2)\tilde{g}. \tag{3.5}$$

A. Real-space path integrals

Having seen the technique in this simple case, the generalization to d dimensions is straightforward. For solutions localized near $\boldsymbol{x} = \boldsymbol{x}'$

$$\frac{\partial}{\partial \Delta t}\tilde{g} = \sum_{n,k}\left(\frac{1}{n!}\prod_m [i\boldsymbol{q}\cdot\mathbf{b}^{(n)}(\boldsymbol{k})]^{n_m}\right)\tilde{g}\,. \tag{3.6}$$

The solution is given in the limit of $\Delta t \to 0$, by

$$\tilde{g} = \prod_{n,k}\tilde{g}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{x}')$$
$$= \prod_{n,k}\exp\left[\frac{\Delta t}{n!}\prod_m [i\boldsymbol{q}\cdot\mathbf{b}^{(n)}(\boldsymbol{k})]^{n_m}\right]. \tag{3.7}$$

By convolving successive Green's functions, one can obtain in principle a path-integral solution [45–48] in the original phase-space, thus using a technique that is well known for diffusion problems. Here we employ the simplest method of integrating over initial values which gives an Ito type of stochastic calculus [39], rather than more sophisticated techniques that are used to obtain path integrals invariant under coordinate transformations [48].

One can verify on differentiation that the finite-time solution in the original space after time $t = N\Delta t$ is a path integral over a set of multidimensional paths $\mathbf{x}(t)$, which are discretized as $(\mathbf{x}^{(0)},\ldots\mathbf{x}^{(N)})$ with $\mathbf{x}(t) = \mathbf{x}^{(N)}$, and

$$d[\mathbf{x}] = \prod_{\ell=0}^{N-1} d\mathbf{x}^{(\ell)}, \tag{3.8}$$

so that, for an initial condition $f(\mathbf{x}^{(0)})$,

$$f(\mathbf{x},t) = \lim_{\Delta t\to 0}\int d[\mathbf{x}]\left(\prod_{\ell=0}^{N-1} g(\mathbf{x}^{(\ell+1)},\Delta t|\mathbf{x}^{(\ell)})\right) f(\mathbf{x}^{(0)})\,. \tag{3.9}$$

To sample this multidimensional integral using probabilistic sampling is not generally possible in the original space, irrespective of whether this limit actually exists or not. The Green's functions inside the integral, $g(\mathbf{x},\Delta t|\mathbf{x}')$, cannot in general have a positive, probabilistic solution owing to Pawula's theorem [49]. Accordingly, $\mathbf{x}(t)$ is not a stochastic path in the original d-dimensional space, since g is not generally positive. Other methods can be used to define higher-order paths [20,21], but these are not well understood in arbitrary space dimensions.

B. Complex space path-integrals

The alternative that is utilized in this paper is to define an equivalent complex space path-integral in a $2d$ (real) dimensional space coordinate $\mathbf{z}$, with volume measure $d\mathbf{z}$. If the original space had d real dimensions, it becomes complex with d complex dimensions. If it was originally complex, it is transformed to twice the original complex dimension. A complex *valued* path integral may be used as another alternative, although it is the positive, probabilistic case that is treated here.

The path-integral equivalence is based on having equal characteristic functions, so consequently, the doubled space Green's functions must have the same Fourier transforms in $\mathbf{q}$ as the original Green's functions. In the Wirtinger case, where the space is originally complex, the two quasi-independent Wirtinger variables $[\boldsymbol{x},\boldsymbol{x}^*]$ are mapped into two independent complex paths, treated as single complex vector $\boldsymbol{z}$ of twice the original complex dimension.

The finite-time complex path-integral solution is constructed using a positive initial distribution $F_0(\mathbf{z}^{(0)})$, whose moments are equal to the KME initial condition, $f_0(\mathbf{x})$:

$$F(\mathbf{z},t) = \int\prod_{\ell=0}^{N-1}[d\mathbf{z}^{(\ell)}G(\mathbf{z}^{(\ell+1)},\Delta t|\mathbf{z}^{(\ell)})]F_0(\mathbf{z}^{(0)})\,. \tag{3.10}$$

The Fourier transform at each step is then

$$\tilde{G}(\mathbf{q},\Delta t|\mathbf{z}') \equiv \int e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{z}')}G(\mathbf{z},\Delta t|\mathbf{z}')d\mathbf{z}\,, \tag{3.11}$$

and this can be factorized to give the form

$$\tilde{G}(\mathbf{q},\Delta t|\mathbf{z}') = \prod_{n,k}\tilde{G}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{z}')\,. \tag{3.12}$$

From Eq. (3.7), the overall multidimensional characteristic Green's function is a product of terms $\tilde{g}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{x}')$. Each of these can now be equated to a complex space Green's characteristic function $\tilde{G}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{z}')$ in d complex dimensions, which is a Fourier transform of a complex propagator:

$$\tilde{G}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{z}') \equiv \int e^{i\mathbf{q}\cdot(\mathbf{z}-\mathbf{z}')}G^{(n,k)}(\mathbf{z},\Delta t|\mathbf{z}')d\mathbf{z}\,. \tag{3.13}$$

Each diffusion coefficient must therefore be analytically continued into the complex plane in the case of variable coefficients in order to have a path-integral solution that equates the two moments at each point in time, while still satisfying the characteristic function time-evolution equations. This generates a positive, dimension-doubled distribution with equivalent analytic moments to the original Wigner distribution, provided the solution converges.

C. Stochastic noise moments

For an HSDE to be equivalent to the Kramers-Moyal equation, the stochastic increments $\Delta\mathbf{z}$ in a finite-time interval must generate the same Green's function after ensemble averaging. From Eq. (3.11), this means that at each step one must prove that, as $\Delta t \to 0$,

$$\tilde{f}(\boldsymbol{q},t+\Delta t) = \int \tilde{G}(\mathbf{q},\Delta t|\mathbf{z}')e^{i\boldsymbol{q}\cdot\boldsymbol{z}'}d\boldsymbol{z}'\,\tilde{f}(\boldsymbol{q},t)$$
$$= \int \tilde{g}(\mathbf{q},\Delta t|\mathbf{x}')e^{i\boldsymbol{q}\cdot\boldsymbol{x}'}d\boldsymbol{x}'\,\tilde{f}(\boldsymbol{q},t). \tag{3.14}$$

From Eq. (3.7), the complex characteristic Green's function $\tilde{G}(\mathbf{q},dt|\mathbf{z}')$ is a product of independent Green's-function terms $\tilde{G}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{z}')$. In Eq. (2.3), the equivalent stochastic process over this interval is postulated to be a sum of terms, one for each (n,k) value in the tensor decomposition, so that $\Delta\mathbf{z}=\sum_{n,k}\Delta\mathbf{z}^{(n,k)}$. A particular value of (n,k) is now treated in order to define the stochastic increment $\Delta\mathbf{z}^{(n,k)}$ whose average Fourier transform will generate the Green's function. The two approaches must generate the same characteristic function. From Eq. (3.14), this means that the characteristic Green's functions must have equal Fourier transforms and hence equal analytic moments in $\boldsymbol{x}'$ and $\boldsymbol{z}'$ respectively. Provided the coefficients are analytic, and the relevant moments exist, this means that for an initial location $\boldsymbol{z}'$:

$$\tilde{G}^{(n,k)}(\mathbf{q},\Delta t|\mathbf{z}')=\exp\left[\frac{\Delta t}{n!}\prod_m(i\mathbf{q}\cdot\mathbf{b}(\boldsymbol{k},z'))^{n_m}\right]$$
$$\equiv\langle\exp[i\mathbf{q}\cdot\Delta\mathbf{z}^{(n,k)}]\rangle\,. \tag{3.15}$$

If each $\Delta\mathbf{z}^{(n,k)}$ is independent, then adding many stochastic increments with different (n,k) values will generate a product of Green's functions, as required. The proof then proceeds as follows. Both the Green's function and the expectation value of the stochastic increment are expanded as multinomials. Next, a noise ansatz is proposed for the stochastic increment, $\Delta\mathbf{z}^{(n,k)}$. Finally, this is inserted in the multinomial expansion so that it can be checked to correspond to the required Green's-function value.

As the first step, the Green's function after expansion has the form

$$\tilde{G}^{(n,k)}(\mathbf{q},\Delta t)=\sum_{p=0}^{\infty}\frac{1}{p!}\left[\frac{\Delta t}{n!}\right]^p\prod_m(i\mathbf{q}\cdot\mathbf{b}(\boldsymbol{k},z'))^{pn_m}\,. \tag{3.16}$$

Here p is an integer power from the expansion of an exponential. The vector index $\boldsymbol{k}$ is defined as $\boldsymbol{k}\equiv(k,m)$, where k identifies an independent additive term in the tensor expansion of $D^{(n)}_{\mathbf{j}}$, while m indexes multiplicative terms occurring in the *same* tensor product. For ease of notation, let $\mathbf{n}=\mathbf{n}(k)$, so n_m is the power of each term. From now on, the indices (n,k) are understood implicitly where it helps to simplify the notation.

The stochastic increment $\Delta\mathbf{z}=\Delta\mathbf{z}^{(n,k)}$ is now assumed to have the following properties as $\Delta t\to 0$, which will be shown to give the required Green's function:

(1)

$$\Delta\mathbf{z}^{(n,k)}=\sum_m\Delta\mathbf{z}^{(n,k,m)}=\sum_m\mathbf{b}(\boldsymbol{k},z')\Delta w^{(\mathbf{n}_k)}_{k,m},$$

(2)

$$\left\langle\prod_m\left[\Delta w^{(\mathbf{n}_k)}_{\boldsymbol{k}}\right]^{pn_m}\right\rangle=\frac{\Delta t^p}{(n!)^p\,p!}\prod_{m=1}(pn_m)!\,,$$

(3)

$$\left\langle\prod_m\left[\Delta w^{(\mathbf{n}_k)}_{\boldsymbol{k}}\right]^{P_m}\right\rangle=0\ if\ P_m\neq pn_m.$$

Defining $G_s\equiv\langle\exp[i\mathbf{q}\cdot\Delta\mathbf{z}^{(n,k)}]\rangle=\langle\prod_m\exp[i\mathbf{q}\cdot\Delta\mathbf{z}(n,k,m)]\rangle$ and exchanging orders of product and sum, where $P=P_1\ldots P_M$ is a vector of powers in the M exponential expansions, one has

$$G_s=\sum_{\boldsymbol{p}}\left\langle\prod_m\frac{[i\mathbf{q}\cdot\Delta\mathbf{z}(\boldsymbol{k})]^{Pm}}{P_m!}\right\rangle. \tag{3.17}$$

Next, expanding using the first ansatz given above gives

$$[i\mathbf{q}\cdot\Delta\mathbf{z}(\boldsymbol{k})]^{P_m}=[i\mathbf{q}\cdot\mathbf{b}(\boldsymbol{k},z')]^{P_m}\times\left[\Delta w^{(\mathbf{n}_k)}_{k,m}\right]^{P_m}. \tag{3.18}$$

This allows a separation of the stochastic and nonstochastic elements, giving

$$G_s=\sum_{\boldsymbol{p}}\left\{\left[\prod_m\frac{[i\mathbf{q}\cdot\mathbf{b}(\boldsymbol{k},z')]^{P_m}}{P_m!}\right]\left\langle\prod_m\left[\Delta w^{(\mathbf{n}_k)}_{\boldsymbol{k}}\right]^{P_m}\right\rangle\right\}. \tag{3.19}$$

The stochastic expectation value of the ansatz is now evaluated to determine if it matches the required solution for a characteristic function. From the noise ansatz, only indices such that $P_m=pn_m$ can give a nonzero moment. The result is therefore a sum over integer powers p such that

$$G_s=\sum_p\left\{\left[\prod_m\frac{[i\mathbf{q}\cdot\mathbf{b}(\boldsymbol{k})]^{pn_m}}{(pn_m)!}\right]\times\frac{\Delta t^p}{(n!)^p\,p!}\prod_m(pn_m)!\right\}$$
$$=\sum_p\frac{\Delta t^p}{(n!)^p\,p!}\prod_m[i\mathbf{q}\cdot\mathbf{b}(\boldsymbol{k})]^{pn_m}. \tag{3.20}$$

This shows that $G_s=\tilde{G}^{(n,k)}(\mathbf{q},\Delta t)$, as required. The higher-order stochastic noise generates the required characteristic function after averaging, provided the relevant stochastic moment equations are satisfied, and the analytic moments of the distributions are well-defined.

D. Simulating Kramers-Moyal equations

Returning to the original KME problem of Eq. (2.1), the results in the previous section show that the stochastic moments do not depend on the type of tensor factorization for an infinite averaging ensemble. The same third-order diffusion could generate either a scalar $\boldsymbol{n}=(3)$ noise, a binary $\boldsymbol{n}=(2,1)$ noise, or a ternary $\boldsymbol{n}=(1,1,1)$ noise vector, all with the same $\mathbf{b}$ vectors. Provided each noise is generated with the correct moments, the final ensemble average is independent of the choice of tensor decomposition.

Nevertheless, for any finite ensemble, there will be nonzero sampling errors, which are proportional to the distribution variance. As a result, it is always more efficient to choose the noise that generates the most compact distribution. This optimization problem is closely related to the general problem of determining a tensor rank, which also gives a minimal decomposition. Compactness issues are examined next, in Sec. IV.

Assuming that the expansion coefficients $D^{(\mathbf{n})}(\mathbf{x})$ are all polynomials in $\mathbf{x}$, then the problem of propagating a solution to this equation reduces to a stochastic difference equation that iterates a series of finite steps in time. For example, in the case of a one-dimensional equation, one has

$$dz=\sum_{\mathbf{n}>0}[D^{(\mathbf{n})}(\mathbf{z})]^{1/n}dw^{(\mathbf{n})}\,. \tag{3.21}$$

With third-order noise in one space dimension, this reduces to

$$dz = D^{(1)}(z)dt + [D^{(2)}(z)]^{1/2}dw^{(2)} + [D^{(3)}(z)]^{1/3}dw^{(3)}\,. \tag{3.22}$$

If the equations only have up to second-order terms, the Kramers-Moyal equation is a generalized diffusion equation. The above result then replicates an Ito stochastic equation, including dimension doubling for cases of nonpositive-definite diffusion. Dimension doubling can require further regularization even with second-order noise due to power-law tails that may occur in extended spaces, as described elsewhere [42].

Higher-order noise terms develop a larger variance as $\Delta t \to 0$. On the other hand, lower-order terms involve more nonlinearities. Given this situation, one useful strategy may be to use a hierarchy of step sizes. If the lowest-order terms are integrated with smaller step sizes than the higher-order terms, this may reduce sampling error while still permitting a reduced discretization error for the nonlinear terms.

As a result, while these methods are exact for constant coefficients, the algorithms may need optimization for nonlinear coefficients. An additional issue that occurs in these cases is that the small time step limit of $\Delta t \to 0$ is generally accompanied by an increased sampling error, so that these limits do not commute. However, our purpose here is to understand the fundamental properties of higher-order complex noise in this general case.

The numerical examples given later show no convergence problems over the range of step sizes used.

IV. HIGHER-ORDER NOISE GENERATION

In order to simulate these equations, one must generate noises with the correct correlations. In this section, higher-order noises of arbitrary vector order are shown to always exist. From Eq. (2.4), each higher-order noise $\Delta \boldsymbol{w}^{(\mathbf{n})}$ is a vector of M terms, with a corresponding order vector $\mathbf{n}$, such that the total order is

$$n = \sum_{m=1}^{M} \boldsymbol{n}_m\,. \tag{4.1}$$

The noise increment can be written in terms of a dimensionless random variate $\boldsymbol{\zeta}^{(\boldsymbol{n})}$, as

$$\Delta \boldsymbol{w}^{(\mathbf{n})} = \boldsymbol{\zeta}^{(\boldsymbol{n})}\left[\frac{\Delta t}{n!}\right]^{1/n}. \tag{4.2}$$

The higher-order noise variate $\boldsymbol{\zeta}^{(\boldsymbol{n})}$ is defined as having nonvanishing moments of

$$\left\langle \prod_m \left[\zeta_m^{(\mathbf{n})}\right]^{pn_m}\right\rangle = \frac{1}{p!}\prod_m (pn_m)!\,. \tag{4.3}$$

After scaling, these higher-order noise moments depend only on the order vector $\mathbf{n}$.

For order $n > 2$ the higher-order noises are nonunique. Two different approaches are given here for multicomponent higher-order noise: polar phase angle expansions and powers of Gaussians. These methods provide direct ways to generate the required random variables. However, noise of any composite order is always obtainable from multicomponent first-order noise, although this may not be optimal.

A method of generating nth-order stochastic processes using polar coordinates is obtained first. This requires random variates with a γ distribution. An alternative is to use Gaussian distributed variates. These may be faster to generate, as Gaussian random variates have very efficient algorithms.

This section focuses on the case of multicomponent first-order noise $\boldsymbol{\zeta}^{(\mathbf{n}_1)}$ with $\boldsymbol{n}_1 = [1, \ldots 1]$ so each component is of first order, but $M = n > 1$. The overall order n is the sum of the component orders, so this case includes higher-order noise of arbitrary order. A more general vector noise of arbitrary order is treated in the Appendix.

A. Multimode polar noise

The noise terms can be transformed into a random polar angle $\theta_m^{(\boldsymbol{n})}$ and a radial noise $r_m^{(\boldsymbol{n})}$, so that

$$\zeta_m^{(\boldsymbol{n})} = r_m^{(\boldsymbol{n})} \exp\left(i\theta_m^{(\boldsymbol{n})}\right). \tag{4.4}$$

1. Angular noise

The random angles must have the property that for integer powers $\mathbf{P} = (P_1 \ldots P_m)$:

$$\left\langle \prod_{m=1}^{M} \exp\left(i\,P_m\theta_m^{(\mathbf{n})}\right)\right\rangle = \sum_p \delta_{\mathbf{P},p\boldsymbol{n}_1}\,. \tag{4.5}$$

The average is unity if $\mathbf{P} = p\mathbf{n}_1$ for some power p, and zero otherwise.

For the case of $\boldsymbol{n}_1 = [1, \ldots 1]$ the angular variables θ_m are generated from $n-1$ uniform random variates, $0 \leqslant u_m < 1$:

$$\theta_m = 2\pi u_m\,. \tag{4.6}$$

The last term with $m = n$ is then given by

$$\theta_n = -\sum_{m=1}^{n-1} 2\pi u_m\,, \tag{4.7}$$

and hence one has

$$\prod_{m=1}^{n}\left[\exp\left(i\theta_m^{(\mathbf{n})}\right)\right]^p = 1\,. \tag{4.8}$$

For any other moments with $\mathbf{P} \neq p\boldsymbol{n}_1$, the expectation value gives a complex circular average over the random noise term, so that

$$\left\langle \prod_{j=1}^{n}\left[\exp\left(i\theta_m^{(\mathbf{n})}\right)\right]^{P_m}\right\rangle = 0\,. \tag{4.9}$$

2. Multicomponent radial noise

The radial moment equations are

$$\left\langle \prod_{m=1}^{n} r_m^p\right\rangle = (p!)^{n-1}\,. \tag{4.10}$$

The γ distribution, $\gamma(\mu;\alpha,\beta) = \beta^\alpha \mu^{\alpha-1}e^{-\beta\mu}/\Gamma(\alpha)$, has nth moments of

$$\langle \mu^n\rangle = \frac{\Gamma(\alpha+n)}{\beta^n\Gamma(\alpha)}. \tag{4.11}$$

It reduces to the exponential distribution for $\alpha = 1$, with moments of $n!/\beta^n$. In this special case the relevant γ distribution has unit shape and rate, which gives the exponential distribution

$$\gamma(\mu_i;1,1) = e^{-\mu_i}\,. \tag{4.12}$$

The radial variables r_m are fractional powers of products of $n-1$ exponentially distributed random variables:

$$r_m = r = \left[\prod_{i=1}^{n-1}\mu_i\right]^{1/n}\,. \tag{4.13}$$

In this case, exponential distribution variates can be efficiently calculated using the definition $\mu = -\ln(u)$, where u is uniformly distributed on $(0,1)$. Recalling the familiar result for exponential distributions, $\langle\mu^p\rangle = p!$, this gives the correct radial moment result, since

$$\left\langle\prod_{m=1}^{M} r_m^p\right\rangle = \left\langle\prod_{i=1}^{n-1}\mu_i^p\right\rangle = (p!)^{n-1}\,. \tag{4.14}$$

This means that for the M-dimensional first-derivative case, there are $2(M-1)$ uniform random variates required, with $M-1$ angular and $M-1$ radial variates. In the $n=2$ case this construction is equivalent to generating a complex Gaussian random variable and its conjugate. In this example, the noise can be written as

$$\zeta_m^{(1,1)} = r_m^{(1,1)}\exp\left(i\theta_m^{(1,1)}\right), \tag{4.15}$$

where ζ_m is a complex Gaussian noise with unit variance and correlations of $\langle|\zeta_{1,2}|^{2p}\rangle = p!$. It has an expansion of $\zeta_{1,2} = (x \pm iy)/\sqrt{2}$. The random variables x, y are independent real Gaussian noises with unit variance, $x = \sqrt{-2\ln(u_1')}\cos(\theta_1)$ and $y = \sqrt{-2\ln(u_1')}\sin(\theta_1)$. This combination of two uniform variates to generate a single complex Gaussian or two real Gaussian random variates is the Box-Muller transformation [50]. This is a known way to generate Gaussian variables in pairs [51], used in numerical calculations.

B. Gaussian product noise

Next, Gaussian variates will be used to treat multimode higher-order stochastic processes as an alternative to polar decomposition. This method may be more convenient in that it uses commonly available Gaussian random variables.

The case with $\mathbf{n} = (1,\ldots 1)$ has nonvanishing indices that correspond to products of single derivatives of different modes. A model for nth-order noise generation is then

$$\zeta_m^{(\mathbf{n}_1)} = \chi_m^*\chi_{m+1}\,[m \leqslant n], \tag{4.16}$$

where $\chi_1 = 1 = \chi_{n+1}$, and for $1 < i,j < n+1$, χ_i are independent complex Gaussian noise variates such that $\langle\chi_i\chi_j^*\rangle = \delta_{ij}$. The lowest-order examples are as follows:

(1) One component: $\zeta^{(1)} = 1$.

(2) Two components: $\zeta_1^{(1,1)} = \chi$, $\zeta_2^{(1,1)} = \chi^*$.

(3) Three components: $\zeta_1^{(1,1,1)} = \chi_2$; $\zeta_2^{(1,1,1)} = \chi_2^*\chi_3$; $\zeta_3^{(1,1,1)} = \chi_3^*$.

To have a nonvanishing moment requires the powers of all noises to be equal, and also that all the noise terms must occur in the overall moment, in order to obtain partnered terms like $\langle[\chi_i^*\chi_i]^p\rangle$ which are nonzero. Hence, the result matches the radial moment requirements of Eq. (4.10), namely,

$$\left\langle\prod_{j=1}^{n}\left[\zeta_m^{(\mathbf{n}_1)}\right]^p\right\rangle = \prod_{j=2}^{n}\langle[\chi_m^*\chi_m]^p\rangle = [p!]^{n-1}\,. \tag{4.17}$$

C. Radial variance

As a point of comparison, it is useful to obtain a radial variance of the noise distribution. This allows an estimate of the sampling error, as more compact noise distributions lead to a lower sampling error. This is highly dependent on the vectorial order of the noise and on the method used to generate the noise. In all cases investigated, the Gaussian and polar methods lead to different radial variances for $n > 2$, with lower radial variance for the polar method.

The radial variance of a given noise component is defined as

$$\sigma^2 = \left\langle\left|\zeta_m^2\right|\right\rangle. \tag{4.18}$$

This is not an analytic moment, since it involves both ζ_m and ζ_m^*. As a consequence, it is not constrained by the original moment equations, which only depend on the analytic moments, and varies with the noise generation algorithm.

1. Polar method

For the scalar noise $z^{(n)}$, this result is known [14], and the radius is a decreasing function of n:

$$\sigma_P^2(n) = \frac{2}{\Gamma(1+2/n)}\,. \tag{4.19}$$

In the case of first-order multicomponent noise with $\boldsymbol{n} = (1,1,\ldots)$, one has that

$$\zeta_m = \left[\prod_{i=2}^{n}\mu_i\right]^{1/n}\exp(i\theta_m)\,. \tag{4.20}$$

Defining the polar transformation variance as $\sigma_P^2(\boldsymbol{n}) = \langle|\zeta_m^2|\rangle$ gives the result that

$$\sigma_P^2(\boldsymbol{n}) = \prod_{i=2}^{n}\langle\mu_i^{2/n}\rangle = [\Gamma(1+2/n)]^{n-1}. \tag{4.21}$$

As an example, consider the strategy of generating nth-order scalar noise from adding up n vector noise components. If this algorithm is used, then since the noise variances add,

$$\sigma_{1P}^2(n) = n[\Gamma(1+2/n)]^{n-1}\,. \tag{4.22}$$

For $n > 2$, this strategy is less compact than using the nth-order scalar noise generated directly with a γ-function method, as described in the Appendix.

2. Gaussian product method

In this case, each Gaussian in the product has an equal variance. Therefore, the Gaussian product generation method

TABLE I. Table of radial variances for polar noise $[\sigma_P^2(\boldsymbol{n})]$ versus Gaussian product (σ_G^2) generation of multidimensional higher-order noise $\boldsymbol{\zeta}^{(n)}$ of order n, with vectorial order $\mathbf{n} = (1, \ldots 1)$. For $n > 2$ the polar noise distribution is more compact, which gives lower sampling errors.

n	2	3	4	5	6
$\sigma_P^2(\boldsymbol{n})$	1	0.8149491	0.696041	0.6197422	0.5678158
σ_G^2	1	1	1	1	1

for higher-order noise gives a variance of

$$\sigma_G^2 = \langle |r_i|^2 \rangle = \langle |\chi_i|^2 \rangle \langle |\chi_{i+1}|^2 \rangle = 1 \,. \tag{4.23}$$

This result is compared with the polar method in Table I. The most compact distribution is obtained using polar noise methods. The table shows that for second-order noise, the distribution radial variances are identical, as expected. However, polar third-order noise gives approximately 20% reduced radial variance, and therefore a lower sampling error. This difference is further increased when going to higher orders.

V. POSITIVE WIGNER FUNCTIONS

The Wigner function [1] has many applications in physics and in other disciplines. It also has a well-known drawback: It is not a positive probability, which led to a famous correspondence involving Dirac, who was concerned about the lack of positivity, and Moyal [52], who developed the dynamical equations for the Wigner function. Indeed, Moyal had hoped that this method would have a direct statistical interpretation, although this is not the case.

A truncated version of the Moyal time-evolution equations was introduced by Graham [53] and has been used in many applications. This is generally positive and restricted to the original phase-space. However, it is not equivalent to quantum mechanics, because the truncated Wigner (TW) representation is an approximate theory where the highest-order derivative terms are truncated by using a $1/N$ expansion for N bosons per mode.

Quantum field dynamical simulations [16,54] have been carried out using this technique. These are in excellent agreement with experimental measurements of quantum polarization squeezing using optical fibers [55], as well as with nontruncated positive-P representation simulations [56,57]. This is due to the large mode occupation numbers used. While TW simulations of Bose-Einstein condensates [16,58–62] can be very useful, with low mode occupations and long time-evolution, unphysical results including vacuum mode depletion can occur [19,28].

There are many other recent applications of Wigner function techniques, including recent work on the theory of optomechanics [63,64], which allows accurate simulation of these nonlinear coupled systems in regimes where quantum effects are large. However, as all such quantum technologies improve, there is increased interest in regimes of even greater nonclassicality where the usual Wigner function is no longer positive.

As a result, it would be useful to have a stochastic technique for the Wigner function *without* truncating higher-order terms. This requires a Wigner function that is complete and positive. The existence of such a positive distribution is known from studies that treated arbitrary s-ordered [65] positive representations [4–7], thus extending earlier work on the positive-P function. These quantum operator representations are all members of more general classes of Gaussian phase-space representations [30].

Here, the properties of the positive Wigner function are summarized. This is a symmetrically ordered phase-space representation that is also a genuine probability for any quantum state. This is only possible due to an extension of its phase space to complex variables.

A. Definition

A positive Wigner distribution always exists on continuing this phase space into a space of double dimension [4], using similar techniques to those that prove a positive-P distribution always exists [9]. The positive Wigner function $P(\mathbf{x})$ is defined by an expansion of the quantum density matrix using a nonorthogonal operator basis:

$$\hat{\rho} = \int P_W(\mathbf{x},\mathbf{p}) \hat{\Lambda}(\mathbf{x},\mathbf{p}) d^2\mathbf{x} d^2\mathbf{p} \,. \tag{5.1}$$

Here, $(\mathbf{x},\mathbf{p}) \equiv \vec{x}$ are independent complex variables. For definiteness, this could be the Wigner function of a bosonic quantum field, but our present arguments only require that there is a set of d pairs of canonical operators $\hat{\mathbf{x}}$, $\hat{\mathbf{p}}$ with commutators:

$$[\hat{x}_i, \hat{p}_m] = i\hbar\delta_{ij}. \tag{5.2}$$

The volume measure on the resulting $2d$-dimensional complex phase-space variables is $d^{2M}\mathbf{x} d^{2M}\mathbf{p} \equiv d\vec{x}$. By contrast, in the case of the usual Wigner function $\mathbf{x},\mathbf{p}$ are real, the volume measure is restricted to a real phase-space, and the distribution is generally not positive valued.

The basis function $\hat{\Lambda}(\mathbf{x},\mathbf{p})$ is a complete operator basis [30]. It is normalized such that

$$\mathrm{Tr}[\hat{\Lambda}(\vec{x})] = 1 \,. \tag{5.3}$$

If $\mathrm{Tr}[\hat{\rho}] = 1$, clearly the distribution is normalized and has the interpretation of a probability density:

$$\int P(\vec{x}) d\vec{x} = 1 \,. \tag{5.4}$$

The operational definition of $P(\vec{x})$ is through its characteristic function

$$\chi(\vec{k}) = \mathrm{Tr}[\hat{\rho} e^{i\vec{k}\cdot\vec{\hat{x}}}] = \int P_W(\vec{x}) e^{i\vec{k}\cdot\vec{x}} d\vec{x}. \tag{5.5}$$

This allows us to obtain the basis element $\hat{\Lambda}(\vec{x})$ through the definition

$$\hat{\Lambda}(\vec{x}) = \frac{1}{(2\pi)^M} \int d\vec{k} e^{i\vec{k}\cdot(\vec{\hat{x}}-\vec{x})} \,. \tag{5.6}$$

In order to understand how to obtain this, raising and lowering operators $\hat{\boldsymbol{a}}^\dagger, \hat{\boldsymbol{a}}$, where $\hat{\boldsymbol{a}} = (\hat{a}^1, \ldots \hat{a}^d)$, can be introduced. The

commutators are

$$[\hat{a}^i,\hat{a}^{\dagger j}] = \delta^{ij},$$
$$[\hat{a}^i,\hat{a}^j] = 0\ . \tag{5.7}$$

If one defines

$$\hat{\mathbf{a}} = \boldsymbol{\lambda}\mathbf{x} + i(2\hbar\boldsymbol{\lambda}^\dagger)^{-1}\mathbf{p}\,, \tag{5.8}$$

where λ is symmetric and invertible, the required conditions are automatically satisfied. As a result, it is always possible to express a Wigner distribution using raising and lowering operators which correspond to a complex phase-space.

B. Complex equations

Kramers-Moyal equations are frequently expressed in terms of vector complex variables $\boldsymbol{\alpha}$ and their conjugates $\boldsymbol{\alpha}^*$. In such cases, the derivation proceeds with essentially no change, except that the derivatives with respect to $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ are treated as independent variables. The complex derivatives are written $(\partial\alpha_1,\ldots\partial\alpha_{2M})$ in the KME equation, where $\alpha_{M+i}\equiv\alpha_i^*$, and the K-M equation can be written as

$$\frac{\partial f}{\partial t} = \sum_n \frac{(-1)^n}{n!}\sum_{\mathbf{j}}\left[\prod_{\mu=1}^{n}\frac{\partial}{\partial\alpha_{j_\mu}}\right]D_{\mathbf{j}}^{(n)}(\mathbf{x})f\ . \tag{5.9}$$

It is convenient to denote the corresponding stochastic complex coordinates as $(\alpha_1,\ldots\alpha_{2M})$, except that $\alpha_{M+i}\neq\alpha_i^*$. With this notation, the above algorithm is directly applicable, with the convention that the coordinate indices i now range from $1,\ldots 2M$:

$$d\boldsymbol{\alpha} = \sum_{n,\boldsymbol{k}}\mathbf{b}^{(n)}(\boldsymbol{k})dw_{\boldsymbol{k}}^{(\boldsymbol{n}_k)}. \tag{5.10}$$

This means that one must map the original complex variables $(\boldsymbol{\alpha},\boldsymbol{\alpha}^*)$ to a dimension-doubled phase space labeled $\vec{\alpha} = (\boldsymbol{\alpha},\boldsymbol{\beta})$, where $(\boldsymbol{\alpha},\boldsymbol{\beta})$ are now two independent complex variables, with equivalences of

$$\left\langle\left\{\hat{a}_1^{n_1}\hat{a}_2^{n_2}\ldots\hat{a}_1^{\dagger\bar{n}_1}\hat{a}_2^{\dagger\bar{n}_2}\ldots\right\}_{SYM}\right\rangle = \left\langle\alpha_1^{n_1}\alpha_2^{n_2}\ldots\beta_1^{\bar{n}_1}\beta_2^{\bar{n}_2}\ldots\right\rangle. \tag{5.11}$$

Here the notation $\{\}_{SYM}$ refers to a symmetrically ordered operator product.

C. Parametric downconversion

As a simple example, consider the problem of nondegenerate parametric downconversion in a cavity. This is an instructive dynamical problem of general interest. It is exactly soluble in a special case treated here, as well as being a simple prototype of more complex quantum wave-mixing problems. The corresponding experimental method is now the standard technique for generating correlated photon pairs in quantum information.

This has been used in Bell inequality violation experiments [24], quantum Einstein-Podolsky-Rosen paradoxes [66,67] and entanglement. Thus, the problem is not without experimental applications. Here, the simple case of a vacuum initial condition is treated. This state must stay rigorously unchanged, as the state is an eigenstate of the Hamiltonian. The vacuum case is therefore an exactly soluble test for third-order noise. Adding other terms or initial conditions is straightforward.

This is a case where the truncation of the third-order term in the Wigner equation is known to predict highly unphysical and paradoxical behavior. The resulting equations are equivalent to stochastic electrodynamics (SED). This theory is a hidden variable theory, sometimes proposed as an alternative to quantum electrodynamics. The equations are known to predict negative occupation numbers and an anomalous third-order correlation.

The truncated result is, of course, incorrect, although this approximation is still common and useful at large occupation numbers. The full quantum equations, or other quantum phase-space methods without truncation, do not predict any vacuum depletion. Instead, they correctly and physically predict that the occupation numbers and correlations remain zero at all times. This is a sensitive test case for theories of higher-order multidimensional noise.

What is the prediction of the time-evolved Wigner distribution using third-order noise methods? How does it compare with the truncated Wigner approximation?

The full quantum problem typically involves a master equation with decoherence, including pumping and loss terms. Here unitary time-evolution is treated using the fundamental three-wave-mixing Hamiltonian. This describes two low-frequency modes, the signal $(\hat{a}_1)$ and idler $(\hat{a}_2)$, driven by a high-frequency pump $(\hat{a}_3)$. In an interaction picture, the dynamics is completely described by the interaction Hamiltonian:

$$\hat{H} = i\hbar g[\hat{a}_1^\dagger\hat{a}_2^\dagger\hat{a}_3 - \hat{a}_1\hat{a}_2\hat{a}_3^\dagger]\,. \tag{5.12}$$

Using an effective time scale of $\tau = gt$, the time evolution of the Wigner function is described by the following Kramers-Moyal equation:

$$\frac{\partial W}{\partial\tau} = \left[-\frac{\partial}{\partial\alpha_1}\alpha_2^*\alpha_3 - \frac{\partial}{\partial\alpha_2}\alpha_1^*\alpha_3 + \frac{\partial}{\partial\alpha_3}\alpha_1\alpha_2 + \frac{1}{4}\frac{\partial^3}{\partial\alpha_1\partial\alpha_2\partial\alpha_3^*} + \text{H.c.}\right]W\,. \tag{5.13}$$

Defining a third-order noise strength $b = [-3/2]^{1/3}$, the corresponding stochastic equations for the positive Wigner function, with third-order noise, exists on a six-dimensional complex or twelve-dimensional real space. The equations are

$$\begin{aligned} d\alpha_1 &= \beta_2\alpha_3 d\tau + b dw_1 \\ d\alpha_2 &= \beta_1\alpha_3 d\tau + b dw_2 \\ d\alpha_3 &= -\alpha_1\alpha_2 d\tau + b d\bar{w}_1 \\ d\beta_1 &= \alpha_2\beta_3 d\tau + b d\bar{w}_2 \\ d\beta_2 &= \alpha_1\beta_3 d\tau + b d\bar{w}_3 \\ d\beta_3 &= -\beta_1\beta_2 d\tau + b dw_3\,. \end{aligned} \tag{5.14}$$

Here, the two independent third-order noise terms are $dw, d\bar{w}$. From inspection of the KME, these noises have effective dimension $M = 3$, with vector order $\mathbf{n} = (1,1,1)$, and overall order $n = 3$. Accordingly, the noise components are labeled according to the reduced effective dimension indices.

Their nonvanishing third-order correlations are given by

$$\langle dw_1 dw_2 dw_3\rangle = d\tau/6\,,$$
$$\langle d\bar{w}_1 d\bar{w}_2 d\bar{w}_3\rangle = d\tau/6\,. \tag{5.15}$$

The initial correlations for a vacuum state simply correspond to half a virtual quantum per mode, as normal for a symmetrically ordered quantum-mechanical Wigner representation, so that

$$\langle \alpha_i \beta_j\rangle = \tfrac{1}{2}\delta_{ij}\,. \tag{5.16}$$

It is common to drop the third-order derivative term, which is equivalent to setting $b = 0$ in the above stochastic equations. This truncated Wigner approximation is valid at large occupation numbers. In the figures, third-order noise and TW results are compared using two noise generators, the Gaussian and polar types. All except the last two figures have 10^6 samples, a step-size of $\Delta t = 0.01$, and the equations are integrated with a fourth-order Runge-Kutta (RK4) method.

All results were repeated with a larger step size of $\Delta t = 0.02$. There was virtually no change in the physical moments, indicating convergence was reached as a function of step size. However, as expected, the distribution radius is reduced, and the issue is treated in Sec. V E.

In order to calculate sampling errors, the calculations are carried out using 100 subensembles of 10^4 trajectories each. The statistics of the trajectories are non-Gaussian, but after the subensemble average over 10^4 terms, the results are nearly Gaussian distributed according to the central-limit theorem. Sampling errors are then obtained using standard Gaussian estimators for the variance in the mean, which gives very robust statistical estimates.

Coding was carried out with two independent programs and algorithms using different languages to reduce the potential for errors. One was hand coded in the Scilab language, and was checked against xSPDE, a public domain stochastic differential equation solver [68]. The xSPDE toolbox with RK4 was used in the graphs given here, as they can give fourth-order convergence in the deterministic terms. This is not crucial, and overall convergence is not fourth order due to the noise terms.

D. Physical moment predictions

The results of Table I give a lower variance, which implies a lower sampling error, for polar noise. This is the preferred method that is used below unless stated otherwise. The situation treated is the coupled nonlinear parametric oscillator in an initial vacuum state. With truncation, the TW equations lead to unphysical predictions, including vacuum states with negative occupations.

The predictions for the physical, analytic moments are given in Figs. 1 and 2 for the symmetrized signal and pump mode occupation numbers. These should both remain equal to 1/2 in a vacuum state, giving occupation number zero, in the exact quantum theory result. In the truncated equations, equivalent to a stochastic electrodynamics theory, these change in time, giving negative pump occupations.

With the high-order noise equations, all occupation numbers are approximately zero as physically required. The residual errors of about ± 0.001 are statistical and within the computed sampling error obtained with 10^6 samples. They can easily be reduced further. The graphs show that the TW approximation has systematic errors of about 100 times the sampling error, predicting an unphysical transfer of quanta from pump to signal, even in the vacuum state.

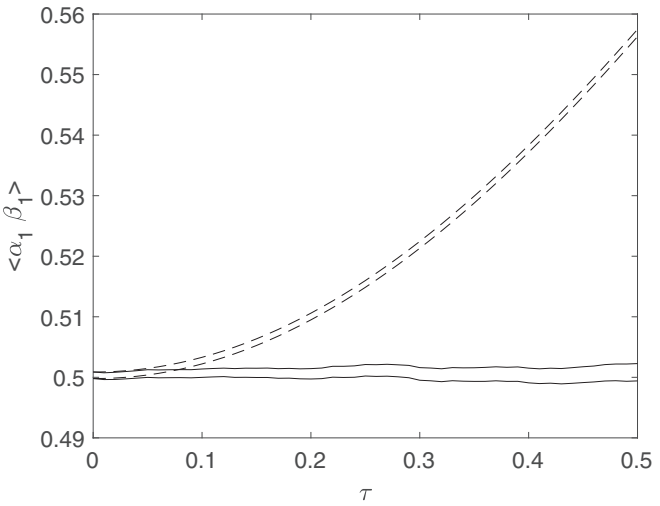

FIG. 1. Time evolution of symmetrized number operator of the parametric amplifier signal mode from a vacuum state, with and without third-order noise terms. The correct Wigner results are $\langle\alpha_1\beta_1\rangle = \langle\hat{n}_1\rangle + 0.5 = 0.5$. Parallel lines indicate sampling error at $\pm\sigma$. Dashed line shows the unphysical increase in apparent number above the vacuum state, without third-order noise. The behavior of $\langle\alpha_2\beta_2\rangle$ is identical. Solid line includes third-order polar noise, showing correctly that the signal occupation number remains approximately zero up to the sampling error (estimated as ± 0.001).

Figure 3 shows how the triple correlation evolves in time. One can readily verify that the expected evolution of the third-order correlations due to higher-order noise is precisely $d\langle\alpha_1\alpha_2\beta_3\rangle = -\frac{1}{4}d\tau$. In this figure the anomalous triple correlation of the TW method, the dashed line, is canceled exactly by the third-order noise so that the overall

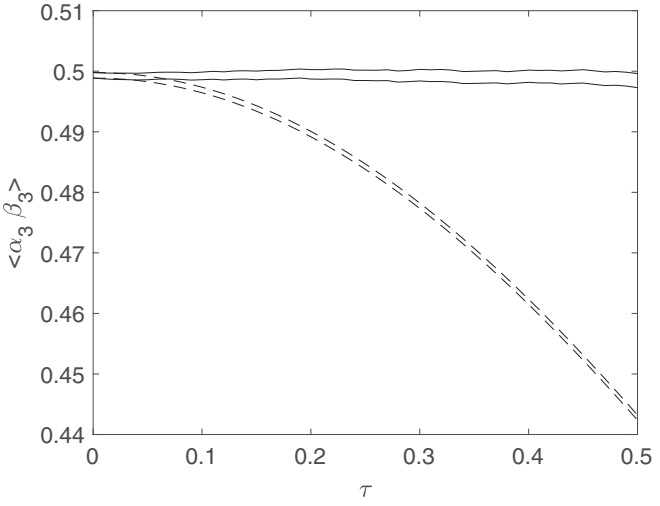

FIG. 2. Time evolution of symmetrized number operator of the pump mode from a vacuum state, with (solid) and without (dashed) third-order noise terms. The correct Wigner results are $\langle\alpha_3\beta_3\rangle = \langle\hat{n}_3\rangle + 0.5 = 0.5$. Other parameters as in Fig. 1.

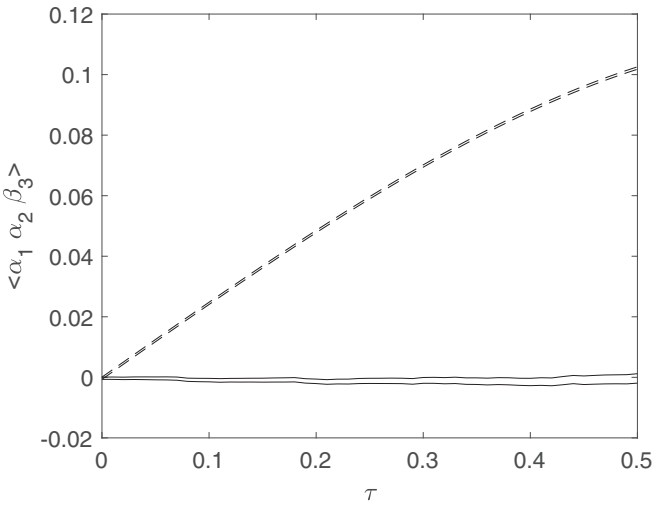

FIG. 3. Time evolution of symmetrized triple correlation of the three modes from a vacuum state, with (solid) and without (dashed) third-order noise terms. Here $\langle\alpha_1\alpha_2\beta_3\rangle = \langle\hat{a}_1\hat{a}_2\hat{a}_3\rangle$. Dashed line shows unphysical increase in the triple correlation, *without* third-order noise. Solid line includes third-order polar noise, showing correctly that the triple correlation is zero up to the sampling error, which gradually increases with increasing distribution radius. Other parameters as in Fig. 1.

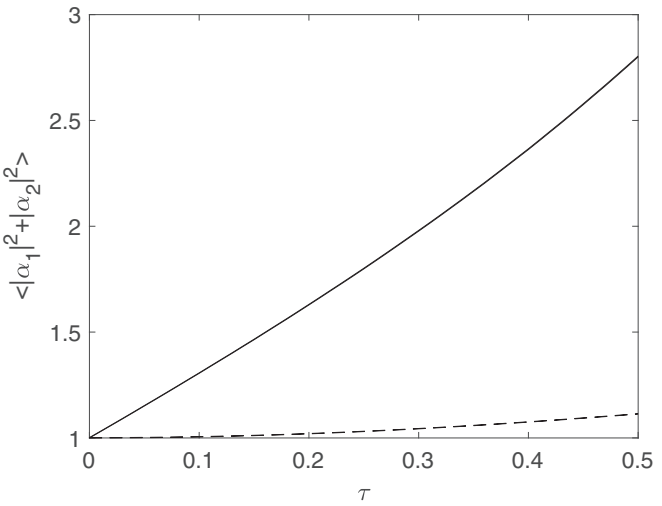

FIG. 4. Time evolution of modulus squared of the signal plus idler phase-space coordinate, with (solid) and without (dashed) third-order polar noise terms. Upper line shows increased modulus with third-order polar noise, lower line gives modulus without third-order noise. The pump phase-space coordinate radius has a similar increase. This shows that the distribution radius increases over time, even when the observables are time invariant. Other parameters as in Fig. 1.

triple correlation remains zero within the sampling error, as expected for a vacuum state.

Analytic predictions are known for the exact and truncated results for the triple correlations, as an expansion in τ, and these agree with the numerical results reported here:

$$\langle\hat{a}_1\hat{a}_2\hat{a}_3\rangle = 0 \ \text{[Exact]},$$
$$\langle\alpha_1\alpha_2\alpha_3^*\rangle = \tfrac{1}{4}\tau + O(\tau^2) \ \text{[TW]}. \qquad (5.17)$$

By $\tau = 1$, the triple-correlation sampling error will start to grow substantially, which is the limit of usefulness of this result without a stabilizing term. This is caused by the general growth in distribution radius. Similar issues occur with the positive-P representation, which is a normally ordered, stochastic representation [69]. Such behavior is inevitable when mapping time-reversible evolution onto a stochastic process with independent paths, since only an inward drift can prevent a growth in radius in a diffusion process. Yet an inward drift would break time-reversal invariance, and it is not present. This issue may be overcome by the use of stochastic gauge stabilizers, but this investigation is outside the scope of the present paper.

This example involves six coupled complex nonlinear stochastic equations in a twelve-dimensional real space, with two independent sets of third-order noise terms. It is therefore a strong test of these methods, and one that is useful in that there is a rigorous, exact solution for all correlations. For the parameters and time scales used here, third-order noise reproduces the known physical results. The results obtained here focus on the question of local noise correlation properties. Optimizing the numerical algorithm for these novel types of noise over longer time scales remains to be investigated.

E. Radial variance

The radial growth of the distribution determines the overall sampling error. The results given above show that the physical moments are all obtained correctly, but it is interesting to investigate the effects of different types of noise generation and step sizes. Results for the growth of the radial variance are given in this section, employing both the polar and Gaussian noise strategies. As expected from the analysis of Table I, polar noise has reduced variance.

Figure 4 uses compact polar noise to show the time evolution of the squared modulus of the signal and idler coordinates and the pump coordinate, with a step size of $\Delta\tau = 0.01$. This is a measure of the distribution size, indicating how the distribution is spreading due to the higher-order noise. This is not an analytic moment. Therefore, while it is is significant in determining the sampling error, it does not correspond to a physical observable of the Wigner function.

The rate of radial growth decreases as the step size increases. This is shown in Fig. 5, which uses larger time steps. However, there is a tradeoff, since increased time steps can also increase discretization errors proportional to $\Delta\tau$. For these results shorter time steps were used to reduce discretization error. There is good convergence in the ensemble sense despite the growth in radial size.

Figures 6 uses Gaussian-generated third-order noise to show the time evolution of the squared modulus of the signal and idler coordinates, again with a step size of $\Delta\tau = 0.01$. This gives a larger variance due to the less compact noise distribution.

Radial growth is more rapid for Gaussian noise than for polar noise. Such larger distribution widths cause an increased sampling error, which is significant when calculating higher-order correlation functions. Gaussian noise can be used, but there is a drawback in that the distributions are less compact.

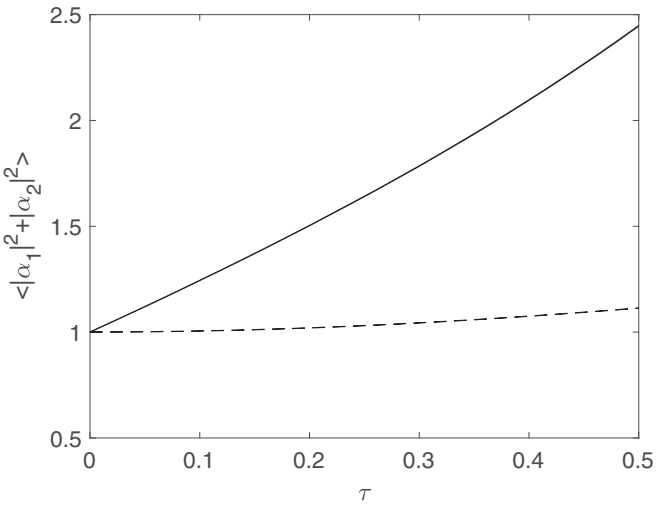

FIG. 5. Time evolution of modulus squared of the signal plus idler radial variance, with (solid) and without (dashed) polar higher-order noise, and with larger time steps of $\Delta\tau = 0.02$. This shows that the stochastic distribution radius increases more slowly with larger time steps. Other parameters as in Fig. 1.

For this reason, the physical moment results were all obtained with polar third-order noise.

VI. OUTLOOK

In summary, this paper has obtained methods for generating stochastic higher-order noise of any vectorial order and in any number of phase-space dimensions, using tensor decomposition theory. Gaussian and polar noise algorithms are derived allowing the efficient generation of any vectorial order of this type of random variate. Due to its greater symmetry, polar noise results in lower variances in all cases treated here. Unlike conventional second-order stochastic noise, the sampling error increases as the step-size is reduced. Despite this, numerical results are in excellent agreement with the predicted moments, and the increases in variance are relatively gradual.

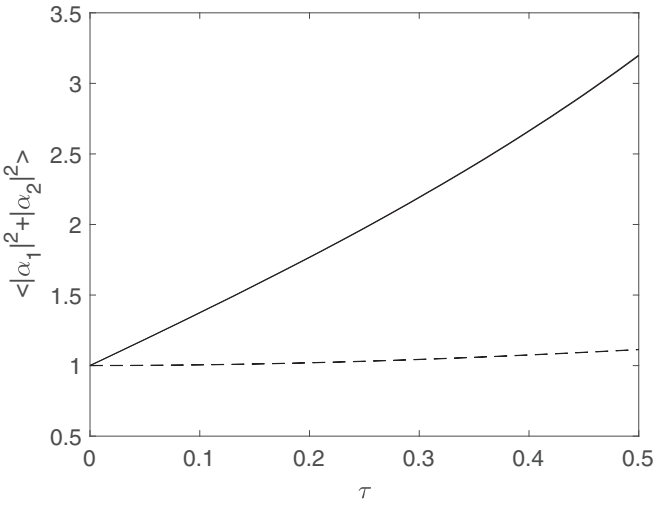

FIG. 6. Time evolution of modulus squared of the signal plus idler radial variance, with (solid) and without (dashed) third-order Gaussian noise terms, and with a step size of $\Delta\tau = 0.01$. This shows that the distribution radius increases over time more quickly for the less compact Gaussian product noise generator. Other parameters as in Fig. 1.

As well as formal derivations of the equations, these results have been calculated explicitly using Green's functions and path integrals, together with convergence properties, methods for generating the stochastic terms, and physical applications of these results. When there are nonlinear terms, such complex equations need to be regularized through the use of weighted trajectories [42], which use the fact that the equations are nonunique.

To test how HSDE techniques can solve vacuum depletion issues, they were applied to an exactly soluble case: the Wigner function for three-mode quantum parametric amplification. This gives rise to six-dimensional higher-noise complex stochastic equations, as each phase-space dimension is doubled. In this case, third-order noise removes unphysical behavior, including negative occupation numbers, found when higher-order terms are truncated from the Wigner equations.

Probabilistic representations of this type are part of a larger class of Gaussian phase-space representations [22,29,30]. These techniques can also involve higher-order derivatives, depending on the exact Hamiltonian or master equation. These terms are particularly an issue with nonlinear master equations, used for representing nonlinear loss in BEC systems [31], as well as in engineered dissipative systems.

HSDE techniques are therefore potentially applicable to a wide variety of nonlinear physical problems. The results obtained here are a preliminary step towards a full understanding of these methods. The main focus of the paper is how to obtain multidimensional HSDE equations from an arbitrary tensor decomposition of a higher-order diffusion matrix. This can give a guide as to when TW methods can fail and provide appropriate correction terms.

ACKNOWLEDGMENT

This work was supported by the Australian Research Council.

APPENDIX: GENERAL HIGHER-ORDER STOCHASTIC NOISE

As explained in Sec. IV, the general noise increment can be written in terms of a dimensionless random variate $\boldsymbol{\zeta}^{(\boldsymbol{n})}$ as

$$\Delta \boldsymbol{w}^{(\mathbf{n})} = \boldsymbol{\zeta}^{(\boldsymbol{n})} \left[\frac{\Delta t}{n!}\right]^{1/n} e^{(n_m \phi_m)}. \tag{A1}$$

where ϕ_m is a scaling factor defined so that $\sum_m \phi_m = 0$. This higher-order noise variate $\boldsymbol{\zeta}^{(\boldsymbol{n})}$ must therefore have nonvanishing moments of

$$\left\langle \prod_m \left[\zeta_m^{(\mathbf{n})}\right]^{pn_m} \right\rangle = \frac{1}{p!} \prod_m (pn_m)!\,. \tag{A2}$$

After scaling, these moments depend only on the order vector $\mathbf{n}$. The main text in Sec. IV explains techniques for generating

these noises for the special case where the vector order is $\mathbf{n} = (1, \ldots 1)$, and the scalar noise generation problem is treated in earlier work [14]. This Appendix gives techniques for generating arbitrary higher-order noises of any vector order.

1. Additive higher-order noise generation

Higher-order noise with $n_m > 1$ is readily obtained from adding n_m multicomponent first-order noise terms of the same total order. Given an nth-order multicomponent first-order noise $\boldsymbol{\zeta}^{(\mathbf{n}_1)}$ with $\boldsymbol{n}_1 = [1, \ldots 1]$ one can generate a random vector $\boldsymbol{\zeta}^{(\mathbf{n})}$ of the same total order on introducing a partial sum, $s_m = \sum_{m'=1}^{m-1} n_{m'}$, and then defining

$$\zeta_m^{(\mathbf{n})} = \sum_{m'=s_m}^{m_1+n_m-1} \zeta_{m'}. \tag{A3}$$

This can be illustrated by constructing a scalar noise $\zeta^{(n)}$ from a first-order vector noise $\boldsymbol{\zeta}^{(\boldsymbol{n}_1)}$. In this case, the moments are

$$\begin{aligned}\langle[\zeta^{(n)}]^{np}\rangle &= \left\langle\left[\sum \zeta_m^{(\boldsymbol{n}_1)}\right]^{np}\right\rangle \\ &= \frac{(np)!}{(p!)^n}\left\langle\left[\prod \zeta_m^{(\boldsymbol{n}_1)}\right]^{p}\right\rangle.\end{aligned} \tag{A4}$$

Here Eq. (4.3) was used, combined with the result that only the most symmetric term in the multinomial expansion with equal powers of ζ_m will be nonzero. Using Eq. (4.3) again for the first-order noise, one can obtain the required result for the nth-order moments, that

$$\langle[\zeta^{(n)}]^{np}\rangle = \frac{(pn)!}{(p!)}. \tag{A5}$$

While this technique is simple, it is relatively inefficient, and the remaining techniques of this Appendix are preferred.

2. Polar higher-order noise generation

A direct generation technique similar to that used for one-dimensional higher-order noise is also possible and is more efficient [14]. The noise term is transformed into a polar angle $\theta_m^{(\boldsymbol{n})}$ and a radial noise $r_m^{(\boldsymbol{n})}$ so that

$$\zeta_m^{(\boldsymbol{n})} = r_m^{(\boldsymbol{n})} \exp\left(i\theta_m^{(\boldsymbol{n})}\right). \tag{A6}$$

From the noise moment equation, the random angles must have the property that for integer powers $\mathbf{P} = (P_1 \ldots P_m)$,

$$\left\langle \prod_{m=1}^{M} \exp\left(i\, P_m \theta_m^{(\mathbf{n})}\right) \right\rangle = \sum_p \delta_{\mathbf{P},p\boldsymbol{n}}\,, \tag{A7}$$

where $\delta_{\mathbf{P},p\boldsymbol{n}}$ is a Kronecker δ function. For noise with only a single nonzero element, so that $M = 1$, the complex noise is generated using the discrete angle technique, already known for the one-dimensional case [14]. In this method, $\theta^{(n)} = 2\pi\ell/n$, where the angle has discrete random values, with integer $\ell = 0, \ldots n-1$.

For $M > 1$, the angular variables θ_m can be generated from at most $M-1$ uniform random variates, $0 \leqslant u_m < 1$:

$$\theta_m = 2\pi u_m\,. \tag{A8}$$

The last term with $m = M$ is then given by

$$\theta_M = -\frac{1}{n_M}\sum_{m=1}^{M-1} 2\pi n_m u_m\,. \tag{A9}$$

With this construction, one has, as required,

$$\prod_{m=1}^{M}\left[\exp\left(i\theta_m^{(\mathbf{n})}\right)\right]^{pn_m} = 1\,. \tag{A10}$$

For any other moments with $\mathbf{P} \neq p\mathbf{n}$, the expectation value gives a complex circular average over the random noise term, so that

$$\left\langle \prod_{j=1}^{M}\left[\exp\left(i\theta_m^{(\mathbf{n})}\right)\right]^{P_m} \right\rangle = 0\,. \tag{A11}$$

To obtain the radial distribution, the moment equations are utilized. From these equations, the radial terms must satisfy

$$\left\langle \prod_{m=1}^{M} r_m^{pn_m} \right\rangle = \frac{1}{p!}\prod_{m=1}^{M}(pn_m)!\,. \tag{A12}$$

To obtain the radial variate, one can define μ_α as a random variable with a gamma distribution $\gamma(\mu_\alpha;\alpha,\beta)$, having equal shape parameter (α) and rate parameter (β), so that $\alpha = \beta$. Alternatively, if ν_α is a random variable with a gamma distribution $\gamma(\nu_\alpha;\alpha,1)$, then $\mu_\alpha = \nu_\alpha/\alpha$. Since the distribution is not uniquely defined by the moment restrictions, it is useful to employ additional constraints. This depends on the requirements for sampling errors in each direction. It is assumed here that it is desirable to have equal radial moments of n_mth order in all noise components, so the choice is made that all terms of form $r_m^{n_m}$ are made equal and have a distribution derived from the γ distribution.

The required distribution is obtained upon extending Eq. (4.14), with the general definition of

$$r_m^{n_m} = \prod_{m=1}^{M}\left(n_m!\prod_{j=1}^{n_m}\mu_{j/n_m}^{(m)}\right)^{1/M}. \tag{A13}$$

Here each variable $\mu_{j/n_m}^{(m)}$ is a random, gamma distributed variable except for $\mu_1^{(M)} \equiv 1$. This includes as special cases:

(1) $M = n$ vector noise, from Eq. (4.13): $r_m^n = \prod_{m=1}^{n-1}\mu_1^{(m)}$

(2) $M = 1$ scalar noise [14]: $r^n = n!\prod_{k=1}^{n-1}\mu_{j/n_1}$

Hence, using the gamma distribution moment result of Eq. (4.11), the radial moments generated with this method can be shown to be

$$\left\langle\left[\prod_{j=1}^{M} r_j^{n_j}\right]^p\right\rangle = \frac{1}{p!}\prod_{m=1}^{M}(n_m!)^p\prod_{k=1}^{n_m}\frac{\Gamma(k/n_m+p)}{(k/n_m)^p\Gamma(k/n_m)}\,. \tag{A14}$$

Since the γ variates are all independent, we can directly apply the product theorem for gamma functions [70]. From this result, one can prove the identity [14]:

$$\prod_{k=1}^{n_m}\Gamma(k/n_m+p) = (2\pi)^{(n_m-1)/2}n_m^{-1/2-n_m p}(n_m p)!\,. \tag{A15}$$

Using this identity, one obtains

$$\prod_{k=1}^{n_m} \frac{\Gamma(k/n_m + p)}{(k/n_m)^p \Gamma(k/n_m)} = \frac{(n_m p)!}{(n_m!)^p}. \tag{A16}$$

Substituting this result into Eq. (A14), we obtain the required moment equation:

$$\left\langle \left[\prod_{m=1}^{M} r_m^{n_m} \right]^p \right\rangle = \frac{1}{p!} \prod_{m=1}^{M} (n_m p)!. \tag{A17}$$

Thus, the multimode γ variate product distribution reproduces the required moments of Eq. (4.3). For example, in the scalar $M = 1, n = 2$ case, $r_m = 2\mu_1$. As a result, $\langle |\zeta^{(2)}|^{2p} \rangle = \langle (2\mu_1)^p \rangle = 2^p p!$. The corresponding differential noise term has $\langle dw^2 \rangle = dt$.

3. General Gaussian product noise

To treat the general case, suppose the higher-order noise is generated from adding Gaussian product variables in a strategy similar to that of Plimak *et al.* [13,71]. This choice is not optimal in terms of sampling error but may be faster or simpler, depending on the relative speed of generating Gaussian and γ random variates. With this approach, the ansatz is

$$\zeta_m^{(\mathbf{n})} = \eta_m^* + \chi_m^* \chi_{m+1} \eta_m^{n_m - 1} \; [m \leqslant M]. \tag{A18}$$

As above, $\chi_1 = 1 = \chi_{M+1}$, while η_i and χ_i are independent complex Gaussian noise variates such that apart from the two end-cases,

$$\langle (\eta_i \eta_j^*)^p \rangle = \langle (\chi_i \chi_j^*)^p \rangle = \delta_{ij} p!. \tag{A19}$$

The χ terms act as a link that chains all the noises together, while the η term is used to generate the type of internal interference that is necessary for higher derivative noise when $n_m > 1$. If $n_m = 1$, one may set $\eta_m = 0$ to regain the single-derivative result of Eq. (4.16).

The general noise moments are now given from the binomial theorem by

$$\left\langle \prod_{m=1}^{M} \left[\zeta_m^{(\mathbf{n})}\right]^{P_m} \right\rangle = \left\langle \prod_{m=1}^{M} \sum_{p_m=0}^{n_m} \frac{\left[\chi_m^* \chi_{m+1} \eta_m^{n_m - 1}\right]^{p_m}}{(p_m)!(P_m - p_m)!} \times P_m! (\eta_m^*)^{P_m - p_m} \right\rangle. \tag{A20}$$

The chainlike nature of the correlations of $\chi_j^* \chi_{j+1}$ mean that the only terms that survive in the overall product of sums have the same power $p_1 = \cdots = p_M = p$. In addition, the correlation properties of complex Gaussian noise η_i mean that the nonzero terms must have $P_m = p n_m$.

There is always a term of this order in each sum, since $P_m \geqslant p$. This implies that only the required moments are nonvanishing. Any moment P_m that is not an integer multiple of n_m will vanish, and the multiplying factor must be the same for each m. Hence, using Gaussian moment properties and recalling that $\chi_1 = 1$, one obtains the correct nonvanishing moments as required, given by

$$\left\langle \prod_{m=1}^{M} \left[\zeta_m^{(\mathbf{n})}\right]^{p n_m} \right\rangle = \frac{1}{p!} \prod_{m=1}^{M} (p n_m)!. \tag{A21}$$

[1] E. Wigner, Phys. Rev. **40**, 749 (1932).
[2] M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, Phys. Rep. **106**, 121 (1984).
[3] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).
[4] F. A. M. de Oliveira, Phys. Rev. A **45**, 3113 (1992).
[5] F. A. M. de Oliveira, Phys. Rev. A **45**, 3350 (1992).
[6] F. A. M. de Oliveira, Phys. Rev. A **45**, 5104 (1992).
[7] S. Chaturvedi, G. S. Agarwal, and V. Srinivasan, J. Phys. A **27**, L39 (1994).
[8] S. Chaturvedi, P. Drummond, and D. F. Walls, J. Phys. A **10**, L187 (1977).
[9] P. D. Drummond and C. W. Gardiner, J. Phys. A **13**, 2353 (1980).
[10] P. Sainty, J. Math. Phys. **33**, 3128 (1992).
[11] M. K. Olsen, L. I. Plimak, and M. J. Collett, Phys. Rev. A **64**, 063601 (2001).
[12] M. K. Olsen, L. I. Plimak, and M. Fleischhauer, Phys. Rev. A **65**, 053806 (2002).
[13] L. I. Plimak, M. Fleischhauer, M. K. Olsen, and M. J. Collett, Phys. Rev. A **67**, 013812 (2003).
[14] P. D. Drummond, J. Phys. A: Math. Theor. **47**, 335001 (2014).
[15] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed. (Springer-Verlag, Berlin, 1985), p. 442.
[16] M. J. Steel, M. K. Olsen, L. I. Plimak, P. D. Drummond, S. M. Tan, M. J. Collett, D. F. Walls, and R. Graham, Phys. Rev. A **58**, 4824 (1998).
[17] P. B. Blakie, A. S. Bradley, M. J. Davis, R. J. Ballagh, and C. W. Gardiner, Adv. Phys. **57**, 363 (2008).
[18] P. Drummond and P. Kinsler, Quantum Semiclassical Opt. **7**, 727 (1995).
[19] P. Deuar and P. D. Drummond, Phys. Rev. Lett. **98**, 120402 (2007).
[20] K. J. Hochberg, Ann. Probab. **6**, 433 (1978).
[21] A. Polkovnikov, Ann. Phys. **325**, 1790 (2010).
[22] J. F. Corney and P. D. Drummond, Phys. Rev. Lett. **93**, 260401 (2004).
[23] D. F. Walls, Nature (London) **306**, 141 (1983).
[24] M. D. Reid and D. F. Walls, Phys. Rev. A **34**, 1260 (1986).
[25] P. Kinsler and P. D. Drummond, Phys. Rev. A **43**, 6194 (1991).
[26] P. Kinsler and P. D. Drummond, Phys. Rev. A **52**, 783 (1995).
[27] P. Kinsler, Phys. Rev. A **53**, 2000 (1996).
[28] A. Sinatra, C. Lobo, and Y. Castin, J. Phys. B **35**, 3599 (2002).
[29] J. F. Corney and P. D. Drummond, J. Phys. A **39**, 269 (2006).
[30] J. F. Corney and P. D. Drummond, Phys. Rev. A **68**, 063822 (2003).
[31] B. Opanchuk and P. D. Drummond, J. Math. Phys. **54**, 042107 (2013).
[32] P. D. Drummond and P. Deuar, J. Opt. B. **5**, S281 (2003).
[33] J. E. Moyal, Mathematical Proc. Cambridge Philos. Soc. **45**, 99 (1949).
[34] A. D. Fokker, Ann. Phys. **348**, 810 (1914).

[35] M. Planck, Sitzber Preuss. Akad. Wiss. Physik. Math. **K1**, 324 (1917).

[36] W. Wirtinger, Math. Ann. **97**, 357 (1927).

[37] F. L. Hitchcock, J. Math. Phys. **6**, 164 (1927).

[38] C. J. Hillar and L.-H. Lim, J. Assoc. Comput. Mach. **60**, 45 (2013).

[39] K. Itō, *On Stochastic Differential Equations*, Vol. 4 (American Mathematical Society, Providence, RI, 1951).

[40] L. Arnold, *Stochastic Differential Equations* (Wiley Interscience, Hoboken, NJ, 1974).

[41] C. W. Gardiner, *Stochastic Methods*, 1st ed. (Springer, Berlin, 1985).

[42] P. Deuar and P. D. Drummond, Phys. Rev. A **66**, 033812 (2002).

[43] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, SIAM J. Matrix Anal. Appl. **30**, 1254 (2008).

[44] N. Wiener, Acta Mathematica **55**, 117 (1930).

[45] L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).

[46] S. Machlup and L. Onsager, Phys. Rev. **91**, 1512 (1953).

[47] R. L. Stratonovich, Sov. Phys. JETP **31**, 1012 (1956).

[48] R. Graham, Z. Phys. B **26**, 397 (1977).

[49] R. Pawula, IEEE Trans. Inf. Theory **13**, 33 (1967).

[50] G. E. Box and M. E. Muller, Ann. Math. Stat. **29**, 610 (1958).

[51] D. E. Knuth, *Seminumerical Algorithms* (Addison Wesley, Reading, MA, 1998).

[52] A. Moyal, *Maverick Mathematician: The Life and Science of J.E. Moyal* (ANU Press, Canberra, 2006).

[53] R. Graham, in *Springer Tracts in Modern Physics* (Springer, Berlin, 1973), pp. 1–97.

[54] P. D. Drummond and A. D. Hardman, Europhys. Lett. **21**, 279 (1993).

[55] J. F. Corney, P. D. Drummond, J. Heersink, V. Josse, G. Leuchs, and U. L. Andersen, Phys. Rev. Lett. **97**, 023606 (2006).

[56] S. J. Carter, P. D. Drummond, M. D. Reid, and R. M. Shelby, Phys. Rev. Lett. **58**, 1841 (1987).

[57] P. D. Drummond, R. M. Shelby, S. R. Friberg, and Y. Yamamoto, Nature (London) **365**, 307 (1993).

[58] A. A. Norrie, R. J. Ballagh, and C. W. Gardiner, Phys. Rev. Lett. **94**, 040401 (2005).

[59] A. A. Norrie, R. J. Ballagh, and C. W. Gardiner, Phys. Rev. A **73**, 043617 (2006).

[60] M. Egorov, R. P. Anderson, V. Ivannikov, B. Opanchuk, P. D. Drummond, B. V. Hall, and A. I. Sidorov, Phys. Rev. A **84**, 021605 (2011).

[61] B. Opanchuk, M. Egorov, S. Hoffmann, A. Sidorov, and P. D. Drummond, Europhys. Lett. **97**, 50003 (2012).

[62] P. D. Drummond and B. Opanchuk, Phys. Rev. A **96**, 043616 (2017).

[63] S. Kiesewetter, Q. Y. He, P. D. Drummond, and M. D. Reid, Phys. Rev. A **90**, 043805 (2014).

[64] S. Kiesewetter, R. Y. Teh, P. D. Drummond, and M. D. Reid, Phys. Rev. Lett. **119**, 023601 (2017).

[65] K. E. Cahill and R. J. Glauber, Phys. Rev. **177**, 1857 (1969).

[66] M. D. Reid and P. D. Drummond, Phys. Rev. Lett. **60**, 2731 (1988).

[67] M. D. Reid, Phys. Rev. A **40**, 913 (1989).

[68] S. Kiesewetter, R. Polkinghorne, B. Opanchuk, and P. D. Drummond, SoftwareX **5**, 12 (2016).

[69] M. R. Dowling, P. D. Drummond, M. J. Davis, and P. Deuar, Phys. Rev. Lett. **94**, 130401 (2005).

[70] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. (Elsevier/Academic Press, Amsterdam, 2007), pp. xlviii+1171, translated from the Russian, Translation edited and with a preface by A. Jeffrey and D. Zwillinger, with one CD-ROM (Windows, Macintosh, and UNIX).

[71] L. I. Plimak, M. J. Collett, and M. K. Olsen, Phys. Rev. A **64**, 063409 (2001).