LECTURE 10 STOCHASTIC PARABOLIC EQUATIONS AND THEIR NUMERICAL METHODS III METHODS OF PROBABILISTIC CHARACTERISTICS AND WIENER CHAOS METHODS

1. Methods of probabilistic characteristics

1.1. **A motivating example.** Consider *stochastic advection-diffusion equation* with periodic boundary condition, written in the Stratonovich form as

$$\begin{aligned} du(t,x) &= \epsilon u_{xx}(t,x)dt + \sigma u_x(t,x)\circ dW(t), \quad t>0,\ x\in(0,2\pi), \\ u(0,x) &= \sin(x), \end{aligned} \tag{1.1}$$

or in the Itô form as

$$du(t,x) = au_{xx}(t,x)\,dt + \sigma u_x(t,x)\,dW(t), \quad u(0,x)=\sin(x).$$

Here $W(t)$ is a standard one-dimensional Wiener process, $\sigma>0$, $\epsilon\geq 0$ are constants, and $a=\epsilon+\sigma^2/2$. The solution of (1.1) is

$$u(t,x) = e^{-\epsilon t}\sin(x+\sigma W(t)), \tag{1.2}$$

and its first and second moments are

$$\mathbb{E}[u(t,x)] = e^{-at}\sin(x), \quad \mathbb{E}[u^2(t,x)] = e^{-2\epsilon t}\left(\frac{1}{2}-\frac{1}{2}e^{-2\sigma^2 t}\cos(2x)\right).$$

The solution of (1.1) with $\epsilon=0$ (the degenerate case) can be represented via the method of characteristics [Rozovsky and Lototsky, 2018]:

$$u(t,x) = \sin(X_{t,x}(0)), \tag{1.3}$$

where $X_{t,x}(s)$, $0\leq s\leq t$, is the solution of the system of backward characteristics

$$dX_{t,x}(s) = \sigma \overleftarrow{dW}(s), \quad X_{t,x}(t)=x. \tag{1.4}$$

The notation "$\overleftarrow{dW}(s)$" means backward Itô integral. It follows from (1.4) that $X_{t,x}(0)$ has the same probability distribution as $x+\sigma\sqrt{t}\zeta$, where ζ is a standard Gaussian random variable (i.e., $\zeta\sim\mathcal{N}(0,1)$).

Numerical methods. Since we are interested only in computing statistical moments, it is assumed that

$$X_{t,x}(0) = x+\sigma\sqrt{t}\zeta. \tag{1.5}$$

Then, we can estimate the second moment $m_2(t,x) := \mathbb{E}[u^2(t,x)]$ as

$$m_2(t,x) \doteq \hat{m}_2(t,x) = \frac{1}{L}\sum_{l=1}^{L}\sin^2(x+\sigma\sqrt{t}\zeta^{(l)}), \tag{1.6}$$

where $\zeta^{(l)}$, $l=1,\dots,L$, are i.i.d. standard Gaussian random variables. The estimate $\hat{m}_2$ for m_2 is unbiased, and, hence, the numerical procedure for finding m_2 based on (1.6) has only

Date: November 19, 2019.

the Monte Carlo (i.e., statistical) error which can be quantified via half of the length of the 95% confidence interval:

$$\rho_{MC}(t,x) = 2\frac{\sqrt{\mathrm{Var}(\sin^2(x+\sigma\sqrt{t}\zeta))}}{\sqrt{L}} .$$

The statistical error for $\hat{m}_2(t,x)$ from (1.6) (no space-time discretization error in this algorithm), which is computed as

$$2\cdot\max_j \frac{\sqrt{\frac{1}{L}\sum_{l=1}^{L}\sin^4(x_j+\sigma\sqrt{t}\zeta^{(l)}) - [\hat{m}_2(t,x_j)]^2}}{\sqrt{L}}, \tag{1.7}$$

where the set of x_j is the Fourier collocation points and $\zeta^{(l)}$ are as in (1.6).

Exercise 1.1. *Consider the following advection-diffusion equation*

$$du(t,x) = au_{xx}(t,x)dt + \sigma u_x(t,x)dw(t), t>0, x\in\mathbb{R}, 2a>\sigma^2>0. \tag{1.8}$$

Then

$$u(t,x) = v(t,x+\sigma w(t)), v(t,x) = \frac{1}{\sqrt{4c\pi t}}\int_{\mathbb{R}} e^{-|x-y|^2/(4ct)}u(0,y)dy. \tag{1.9}$$

Show that

$$u(t,x) = \mathbb{E}\left(u(0,x+\sqrt{2a-\sigma^2}B(t)+\sigma w(t))|\mathcal{F}_t^w\right)$$

where B is a Wiener process independent of w.

Example 1.2. *Consider the following diffusion-reaction equation*

$$du(t,x) = au_{xx}(t,x)dt + \sigma u(t,x)dw(t), t>0, x\in\mathbb{R}. \tag{1.10}$$

Assume that a,σ are real numbers and $a>0$. Then

$$u(t,x) = u^{heat}(t,x)e^{\sigma w(t)-\left(\sigma^2 t/2\right)}, \tag{1.11}$$

and thus

$$u(t,x) = e^{\sigma w(t)-\left(\sigma^2 t/2\right)}\mathbb{E}u(0,x+\sqrt{2a}B(t)), \tag{1.12}$$

where B is a Wiener process independent of w.

1.2. **Linear parabolic equations with variable coefficients.** Consider the following SPDE written in Itô's form (backward integral):

$$\begin{aligned} du(t,x) &= [\mathcal{L}u(t,x)+f(x)]\,dt + \sum_{k=1}^{q}[\mathcal{M}_k u(t,x)+g_k(x)]*dw_k(t),\ (t,x)\in(0,T]\times\mathcal{D}, \\ u(T,x) &= \psi(x), \quad x\in\mathcal{D}, \end{aligned} \tag{1.13}$$

where

$$\begin{aligned} \mathcal{L}u(t,x) &= \sum_{i,j=1}^{d} a_{ij}(x)\,D_iD_ju(t,x) + \sum_{i=1}^{d} b_i(x)D_iu(t,x) + c(x)\,u(t,x), \\ \mathcal{M}_k u(t,x) &= \sum_{i=1}^{d}\alpha_k^i(x)D_iu(t,x) + \beta_k(x)\,u(t,x), \end{aligned} \tag{1.14}$$

and $D_i := \partial_{x_i}$ and $\mathcal{D}\in\mathbb{R}^d$ is either a periodic domain or that $\mathcal{D}=\mathbb{R}^d$. In the former case we will consider periodic boundary conditions and in the latter the Cauchy problem. Let $(w(t),\mathcal{F}_t) = (\{w_k(t), k\geq 1\}, \mathcal{F}_t)$ be a system of one-dimensional independent standard Wiener processes defined on a complete probability space $(\Omega,\mathcal{F},\mathcal{P})$, where $\mathcal{F}_t,\ 0\leq t\leq T$, is a filtration satisfying the usual hypotheses.

Assumption 1.3.
- *The coefficients of operators $\mathcal{L}$ and $\mathcal{M}$ and their derivatives are uniformly bounded and sufficiently smooth.*
- *(the coercivity condition) $-\mathcal{L}+\frac{1}{2}\sum_{k\geq 1}\mathcal{M}_k\mathcal{M}_k$ is nonnegative definite.*
- *(Smoothness of ψ and polynomial growth) The function $\psi(x)$ is sufficiently smooth and that $\psi(x)$ and its derivatives up to some order belong to the class functions satisfying an inequality of the form*
$$|\psi(x)|\leq K(1+|x|)^{\kappa}$$
where K and κ are positive constants

Denote $\sigma\in\mathbb{R}^{d\times p}$ which satisfies
$$\sigma(x)\sigma^{\top}(x)=2a(x)-\alpha(x)\alpha^{\top}(x). \tag{1.15}$$
Here p is the rank of the left hand side and σ can be computed using Cholesky decomposition.

One probabilistic representation of the solution ([Rozovsky and Lototsky, 2018]) is
$$u(t,x)=E^{w}\left[\varphi\left(X_{t,x}(T)\right)Y_{t,x,1}(T)+Z_{t,x,1,0}(T)\right],\quad T_0\leq t\leq T \tag{1.16}$$
where $X_{t,x}(s)$, $Y_{t,x,y}(s)$, $Z_{t,x,y,z}(s)$, $t\leq s\leq T$, is the solution of the SDEs
$$\begin{aligned} dX &= \left[b(s,X)-\sum_{r=1}^{q}\alpha_r(s,X)\beta_r(s,X)\right]ds && (1.17)\\ &\quad+\sum_{r=1}^{p}\sigma_r(s,X)dW_r(s)+\sum_{r=1}^{q}\alpha_r(s,X)dw_r(s),\quad X(t)=x && (1.18)\\ dY &= c(s,X)Y\,ds+\sum_{r=1}^{q}\beta_r(s,X)Y\,dw_r(s),\quad Y(t)=y && (1.19)\\ dZ &= f(s,X)Y\,ds+\sum_{r=1}^{q}g_r(s,X)Y\,dw_r(s),\quad Z(t)=z && (1.20)\end{aligned}$$
Here $W(s)=(W_1(s),\ldots,W_p(s))^{\top}$ is p-dimensional Wiener process independent of $w(s)$ and $\mathbb{E}^{w}[\cdot]$ means the conditional expectation $\mathbb{E}[\cdot|w(s)-w(t),t\leq s\leq T]$.

Let's consider Euler scheme, for $k=0,\ldots,N-1$,
$$\begin{aligned} \bar{X}_{k+1} &= \bar{X}_k+h\left[b\left(t_k,\bar{X}_k\right)-\sum_{r=1}^{q}\alpha_r\left(t_k,\bar{X}_k\right)\beta_r\left(t_k,\bar{X}_k\right)\right] && (1.21)\\ &\quad+\sum_{r=1}^{q}\alpha_r\left(t_k,\bar{X}_k\right)\Delta_k w_r+\sum_{r=1}^{p}\sigma_r\left(t_k,\bar{X}_k\right)\Delta_k W_r,\quad \bar{X}_0=x, && (1.22)\\ \bar{Y}_{k+1} &= \bar{Y}_k+hc\left(t_k,\bar{X}_k\right)\bar{Y}_k+\sum_{r=1}^{q}\beta_r\left(t_k,\bar{X}_k\right)\bar{Y}_k\Delta_k w_r,\quad \bar{Y}_0=1, && (1.23)\\ \bar{Z}_{k+1} &= \bar{Z}_k+hf\left(t_k,\bar{X}_k\right)\bar{Y}_k+\sum_{r=1}^{q}g_r\left(t_k,\bar{X}_k\right)\bar{Y}_k\Delta_k w_r,\quad \bar{Z}_0=0, && (1.24)\end{aligned}$$
where
$$\Delta_k w_r:=w_r\left(t_{k+1}\right)-w_r\left(t_k\right),\Delta_k W_r:=W_r\left(t_{k+1}\right)-W_r\left(t_k\right) \tag{1.25}$$
Then the numerical solution can be represented by
$$\bar{u}(t,x)=\mathbb{E}^{w}[\psi(\bar{X}_N)\bar{Y}_N+\bar{Z}_N].$$

The Euler method converges with a strong order 1/2:

$$\mathbb{E}[|\bar{u}(t,x) - u(t,x)|^{2p}] \leq C(x)(\Delta t)^p.$$

Also if $\xi_{r,k}$'s (for W_r only)are replaced with discrete random walk ($\mathbb{P}(\zeta = \pm 1) = 1/2$), the convergence order is still 1/2.

1.2.1. *Variance reduction.* To reduce the variance of the term in the conditional expectation, we may use another probabilistic representation of the solution [Milstein and Tretyakov, 2009]:

$$u(t,x) = \mathbb{E}^w[\varphi\left(X_{t,x}(T)\right) Y_{t,x,1}(T) + Z_{t,x,1,0}(T)], \quad T_0 \leq t \leq T \tag{1.26}$$

where $X_{t,x}(s)$, $Y_{t,x,y}(s)$, $Z_{t,x,y,z}(s)$, $t \leq s \leq T$, is the solution of the SDEs

$$dX = \left[b(s,X) - \sum_{r=1}^{q} \alpha_r(s,X)\beta_r(s,X) - \sum_{r=1}^{p} \sigma_r(t,X)\mu_r(s,X)\right] ds \tag{1.27}$$

$$+\sum_{r=1}^{p} \sigma_r(s,X)dW_r(s) + \sum_{r=1}^{q} \alpha_r(s,X)dw_r(s), \quad X(t) = x \tag{1.28}$$

$$dY = c(s,X)Y ds + \sum_{r=1}^{p} \mu_r(s,X)Y\, dW_r + \sum_{r=1}^{q} \beta_r(s,X)Y dw_r(s), \quad Y(t) = y \tag{1.29}$$

$$dZ = f(s,X)Y ds + \sum_{r=1}^{p} \lambda_r(s,X)Y\, dW_r + \sum_{r=1}^{q} g_r(s,X)Y dw_r(s), \quad Z(t) = z \tag{1.30}$$

When $\lambda = 0$ and $\mu = 0$, then it becomes the original representation. When $\lambda \neq 0$ and $\mu = 0$, we can readily observe that

$$\mathbb{E}^w[\int_t^T \sum_{r=1}^{p} \lambda_r(s,X(s))Y(s)\, dW_r(s)] = 0$$

and thus the probabilistic formulation is valid.

The goal is to find μ and λ such that the conditional variance of $\psi(X_{t,x}(T))Y_{t,x,1}(T) + Z_{t,x,1,0}(T)$ is small enough. Observe that

$$\begin{aligned}&\mathrm{Var}^w\left[\varphi\left(X_{t,x}(T)\right) Y_{t,x,1}(T) + Z_{t,x,1,0}(T)\right] \\ &= \mathbb{E}^w[\int_t^T Y_{t,x,1}^2(s,X_{t,x}(s))\sigma_r^i(s,X_{t,x}(s)) \sum_{r=1}^{p} \left[\sum_{i=1}^{d} \sigma_r^i(s,X_{t,x}(s))\frac{\partial v}{\partial x^i} + \mu_r(s,X_{t,x}(s))v + \lambda_r\right]^2 ds].\end{aligned}$$

Thus an ideal variance reduction is reached if

$$\sum_{i=1}^{d} \sigma_r^i \frac{\partial u}{\partial x^i} + \mu_r u + \lambda_r = 0, r = 1,\ldots,p. \tag{1.31}$$

This is not possible as we don't know u and its derivatives. But if we know a function $\tilde{u}(s,x)$ which is close to $u(s,x)$, we can take any μ_r and λ_r such that

$$\sum_{i=1}^{d} \sigma_r^i \frac{\partial \tilde{u}}{\partial x^i} + \mu_r \tilde{u} + \lambda_r = 0, r = 1,\ldots,p. \tag{1.32}$$

Remark 1.4. *It is advantageous to use methods of probabilistic characteristics when solutions at only a few points are needed. The computational cost is relatively low, especially variance reduction methods are applied: only the cost of Monte Carlo methods and cost of solving SODEs while there is no grid in physical space in use. However, the methods of characteristics can be expensive as the methods are usually of first-order convergence (and difficult to obtain* *higher-order convergence). Also, when a solution at a large number of grid points is required, the cost can be prohibitively high.*

2. Wiener chaos solution

Example 2.1 (Multiple Ito integral)**.** *Assume that $W(t)$ is a standard Brownian Motion, show that*

$$\frac{1}{n!}\int_0^t \int_0^{t_n} \dots \int_0^{t_2} dW(t_1)\cdots dW(t_n) = t^{n/2} H_n(\frac{W(t)}{\sqrt{t}}).$$

Here H_n is the n-th Hermite polynomial:

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}. \tag{2.1}$$

When $n = 0$, $H_n(x) = 1$ and we use the convention that when $n < 1$ the integral is defined as 1. When $n = 1$, $\int_0^t dW(t_1) = W(t) = t^{1/2} H_1(\frac{W(t)}{\sqrt{t}})$ as $H_1(x) = x$. Then it can be shown by induction that the integrand in the left hand side is in $\mathbb{L}^2_{\mathrm{ad}}(\Omega; L^2([0,t]))$ and thus the multiple integral is indeed an Ito integral and is equal to the right hand side.

When we want to define the integral $\displaystyle\int_0^t f(t)\, dW(t)$ via a spectral representation of Brownian motion instead of using increments of Brownian motion, we have to use the so-called Ito-Wick product (Wick product) and Stratonovich product.

The use of Ito-Wick product relies on two facts: the integrand f can be expressed as Hermite polynomial expansion of some random variables ("random basis"), and the product is well-defined for this basis. The basis and the Ito-Wick product will be shown and defined shortly. Specifically, let $\{\xi_k\}_{k=1}^\infty$ be a sequence of mutually independent standard Gaussian random variables from a spectral representation of Brownian motion and let $\mathcal{F} = \sigma(\xi_k)_{k\geq 1}$. The following Cameron-Martin theorem states that any element in $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ can be represented by a linear combination of elements in the Cameron-Martin basis (2.2).

Theorem 2.2 (Cameron-Martin)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. The collection $\Xi = \{\xi_\alpha\}_{\alpha\in\mathcal{J}}$ is an orthonormal basis in $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$, where ξ_α's are defined as*

$$\xi_\alpha := \prod_\alpha \left(\frac{H_{\alpha_l}(\xi_l)}{\sqrt{\alpha_l!}} \right), \quad \xi_l = \int_0^t m_l(s)\, dW(s), \quad \alpha \in \mathcal{J}, \tag{2.2}$$

where $\{m_l\}$ is a complete orthonormal basis in $L^2([0,t])$ and $\mathcal{J}$ is the collection of of multi-indices $\alpha = (\alpha_l)_{l\geq 1}$ of finite length, i.e.,

$$\mathcal{J} = \left\{ \alpha = (\alpha_l)_{l\geq 1}, \quad \alpha_l \in \{0, 1, 2, \dots\}, \ |\alpha| := \sum_l \alpha_l < \infty \right\}.$$

Any $\eta \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ can be represented as the following Wiener chaos expansion

$$\eta = \sum_{\alpha\in\mathcal{J}} \eta_\alpha \xi_\alpha, \quad \eta_\alpha = \mathbb{E}[\eta\xi_\alpha], \quad \text{and} \quad \mathbb{E}[\eta^2] = \sum_{\alpha\in\mathcal{J}} \eta_\alpha^2. \tag{2.3}$$

The collection Ξ of random variables $\{\xi_\alpha, \alpha \in \mathcal{J}\}$ is called the Cameron-Martin basis. It can be readily shown that $\mathbb{E}[\xi_\alpha \xi_\beta] = 1$ if $\alpha = \beta$ and 0 otherwise. See below some specific examples of the Cameron-Martin basis.

When $f(t)$ is a continuous semi-martingale with respect to the natural filtration of Brownian motion, we have

$$\int_0^t f(t)\, dW(t) = \int_0^t f(t) \diamond \dot{W}(t)\, dt, \tag{2.4}$$

where the definition of Ito-Wick product "$\diamond$" is based on the product for elements of the Cameron-Martin basis: $\xi_\alpha \diamond \xi_\beta = \sqrt{\dfrac{(\alpha+\beta)!}{\alpha!\beta!}}\xi_{\alpha+\beta}$.

The Cameron-Martin theorem (Theorem 2.2) provides a spectral representation of square-integrable stochastic processes defined on the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This representation is also known as Wiener chaos expansion. The Cameron-Martin basis is listed in Table 1 where there is only one Wiener process in (2.2).

TABLE 1. Some elements of the Cameron-Martin basis ξ_α in (2.2).

$\lvert\alpha\rvert$	α	ξ_α
0	$\alpha = (0, 0, \dots)$	1
1	$\alpha = (0, \dots, 0, 1, 0, \dots)$	$H_1(\xi_i) = \xi_i$
2	$\alpha = (0, \dots, 0, 2, 0, \dots)$	$H_2(\xi_i)/\sqrt{2} = \frac{1}{\sqrt{2}}\left(\xi_i^2 - 1\right)$
2	$\alpha = (0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots)$	$H_1(\xi_i)H_1(\xi_j) = \xi_i\xi_j$

We need in practice to truncate the number of random variables, i.e., let the elements of α be zero with large indexes. To be precise, we introduce the following notation: the order of multi-index α:

$$d(\alpha) = \max\left\{l \geq 1 : \alpha_{k,l} > 0 \text{ for some } k \geq 1\right\}.$$

Also, we need to limit the number of $|\alpha|$. We actually define truncated set of multi-indices

$$\mathcal{J}_{\mathsf{N},\mathsf{n}} = \{\alpha \in \mathcal{J} : \ |\alpha| \leq \mathsf{N}, \ \ d(\alpha) \leq \mathsf{n}\}.$$

In this set, there is a finite number of n dimensional random variables and the number is

$$\sum_{i=0}^{\mathsf{N}} \binom{\mathsf{n}+i-1}{i} = \binom{\mathsf{n}+\mathsf{N}}{\mathsf{N}} = \binom{\mathsf{n}+\mathsf{N}}{\mathsf{n}}.$$

In Table 2, we list the elements in a truncated Cameron-Martin basis. More examples of the

TABLE 2. Elements of a truncated Cameron-Martin basis ξ_α for a finite dimensional random space where $\alpha \in \mathcal{J}_{\mathsf{N},\mathsf{n}}$, $\mathsf{N} = 2$ and $\mathsf{n} = 2$.

$\lvert\alpha\rvert$	α	ξ_α
0	$\alpha = (0, 0)$	1
1	$\alpha = (1, 0)$	$H_1(\xi_1) = \xi_1$
1	$\alpha = (0, 1)$	$H_1(\xi_2) = \xi_2$
2	$\alpha = (2, 0)$	$H_2(\xi_1)/\sqrt{2} = \frac{1}{\sqrt{2}}\left(\xi_1^2 - 1\right)$
2	$\alpha = (0, 2)$	$H_2(\xi_2)/\sqrt{2} = \frac{1}{\sqrt{2}}\left(\xi_2^2 - 1\right)$
2	$\alpha = (1, 1)$	$H_1(\xi_1)H_1(\xi_2) = \xi_1\xi_2$

basis can be generated using the representation of the Hermite polynomial (2.1). Here are the first seven Hermite polynomials:

$$\begin{aligned}
H_0(x) &= 1, \quad H_1(x) = x,\\
H_2(x) &= x^2 - 1,\\
H_3(x) &= x^3 - 3x,\\
H_4(x) &= x^4 - 6x^2 + 3,\\
H_5(x) &= x^5 - 10x^3 + 15x,\\
H_6(x) &= x^6 - 15x^4 + 45x^2 - 15.
\end{aligned}$$

The Hermite polynomials can be represented (computed) by the three-term recurrence relation

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x),\, n \geq 1, \quad H_0(x) = 1,\, H_1(x) = x. \tag{2.5}$$

2.1. **Application to SODEs.** Let's consider the stochastic differential equation

$$dX = W(t)X\,dt, \quad 0 < t \leq 1 \quad X(0) = X_0 = 1, \tag{2.6}$$

where $W(t)$ is a standard Brownian motion. We employ the simplest time discretization – forward Euler scheme. For a uniform partition of $[0, 1]$, $t_k = kh$, $1 \leq k \leq N$ and $Nh = 1$. The forward Euler scheme is

$$X_{k+1} = X_k + W(t_k)X_k h, \quad 0 \leq k \leq N - 1. \tag{2.7}$$

he goal here is to numerically compute the mean and variance of the solution or simply $\mathbb{E}[X_N^p]$, $p = 1, 2$.

Here we notice that $W(t_k)$ need further discretization. We recall that there are two ways to approximate $W(t_k)$. The first one is to use increments of Brownian motions (the spectral approximation is left as an exercise) and the forward Euler scheme becomes

$$X_{k+1} = X_k + \sqrt{h}\sum_{i=0}^{k} \xi_i X_k h, \quad 0 \leq k \leq N - 1. \tag{2.8}$$

Here $\xi_0 = 0$. Then we have that the solution can be represented by $X_N(\xi_1, \xi_2, \ldots, \xi_{N-1})$.

Consider now the Wiener chaos method for (2.8). Suppose that $X_{k+1} = \sum_{\alpha\in\mathcal{J}_{\mathsf{N},k}} x_{\alpha,k}\xi_\alpha$ for $1 \leq k \leq N - 1$. We first apply a Galerkin method in random space – multiplying by the Cameron-Martin basis ξ_β, $\beta \in\in \mathcal{J}_{\mathsf{N},k}$ over both sides of (2.8) and taking expectation (integration over the random space). We then have

$$\mathbb{E}[\xi_\beta \sum_{\alpha\in\mathcal{J}_{\mathsf{N},k}} x_{\alpha,k}\xi_\alpha] = \mathbb{E}[\xi_\beta(1 + \sqrt{h}\sum_{i=1}^{k}\xi_i h)) \sum_{\alpha\in\mathcal{J}_{\mathsf{N},k-1}} x_{\alpha,k-1}\xi_\alpha].$$

By the orthonormality of the Cameron-Martin basis, we have

$$x_{\beta,k} = x_{\beta,k-1}1_{\beta_k=0} + h^{3/2}\sum_{i=1}^{k}\sum_{\alpha\in\mathcal{J}_{\mathsf{N},k-1}} x_{\alpha,k-1}\mathbb{E}[\xi_i\xi_\alpha\xi_\beta].$$

This turns the discrete stochastic equation from the forward Euler scheme into a system of deterministic equations of the Wiener chaos expansion coefficients $x_{\beta,k+1}$. This system of deterministic equation of the coefficients is called *propagator* of the discrete stochastic equation (2.8).

To solve for $\xi_{\beta,k+1}$, one needs to find the expectations of the triples $\mathbb{E}[\xi_i\xi_\alpha\xi_\beta]$. Recalling the recurrence relation (2.5) and orthogonality of Hermite polynomials, the triples are zero unless $\alpha = \beta \pm \varepsilon_i$, where ε_i is a multiindex with $|\varepsilon_i| = 1$ and its only nonzero element is the i-th one. Recalling the (2.5) and when $\beta_i = \alpha_i + 1$, we have

$$\mathbb{E}[\xi_i \xi_{\alpha_i} \xi_{\beta_i}] = \mathbb{E}[\xi_i H_{\alpha_i}(\xi_i) H_{\beta_i}(\xi_i)]/\sqrt{\alpha_i!}/\sqrt{\beta_i!} = \mathbb{E}[H_{\alpha_i+1}(\xi_i) H_{\beta_i}(\xi_i)]/\sqrt{\alpha_i!}\sqrt{\beta_i!} = \sqrt{\alpha_i+1}.$$

Then, the triples can be computed as

$$\mathbb{E}[\xi_i \xi_\alpha \xi_\beta] = \sqrt{\alpha_i+1}\mathbf{1}_{\alpha+\varepsilon_i=\beta} + \sqrt{\beta_i+1}\mathbf{1}_{\beta+\varepsilon_i=\alpha}. \tag{2.9}$$

We have a *propagator* ready for implementation: for $\beta \in \mathcal{J}_{\mathsf{N},k}$,

$$x_{\beta,k} = \big(x_{\beta,k-1} + h^{3/2} \sum_{i=1}^{k-1} \big(\sqrt{\beta_i} x_{\beta-\varepsilon_i,k-1} + \sqrt{\beta_i+1} x_{\beta+\varepsilon_i,k-1} \mathbf{1}_{|\beta|\le \mathsf{N}-1}\big)\big)\mathbf{1}_{\beta_k=0} + h^{3/2} x_{\alpha,k-1}\mathbf{1}_{\beta_k=1}\mathbf{1}_{\alpha=(\beta_1,\dots,\beta_{k-1})}.$$

2.2. **Another representation.** Let $h(t)$ be a smooth, compactly supported function on $(0,1)$, then the collection of the random variables

$$\mathcal{E}_h = \exp\left(\int_0^1 h(t)dw(t) - \frac{1}{2}\int_0^1 h^2(t)dt\right), \quad h \in \mathcal{C}_0^\infty((0,1)) \tag{2.10}$$

s everywhere dense in the space of square-integrable random variables that are measurable with respect to the sigma-algebra of $w(t)$, $t \in [0,1]$. Then square-integrable random process $u = u(t,x)$ is uniquely determined by the collection of deterministic functions

$$u_h(t,x) = \mathbb{E}\left(u(t,x)\mathcal{E}_h\right), h \in \mathcal{C}_0^\infty((0,1)) \tag{2.11}$$

For solutions to SPDEs, very function u_h will satisfy a deterministic PDE

$$du(t,x) = au_{xx}(t,x)dt + \sigma u_x(t,x)dw(t), 0 < t < 1, x \in \mathbb{R} \tag{2.12}$$

where w is a standard Brownian motion, a, σ are non-random constants, $a > 0$, and the initial condition u_0 is non-random. Then

$$\frac{\partial u_h(t,x)}{\partial t} = a\frac{\partial^2 u_h(t,x)}{\partial x^2} + \sigma h(t)\frac{\partial u_h(t,x)}{\partial x}, 0 < t < 1, x \in \mathbb{R} \tag{2.13}$$

where the initial condition $u_h = u_0$ for all functions h.

Choosing h from an orthonormal basis $\{m_k\}_{k\ge1}$ of $L^2(0,1)$ gives $\xi_k = \int_0^1 m_k(t)\,dw(t)$, which are i.i.d. standard normal. Recall the generating function of the Hermite polynomials is

$$e^{xt-\frac{1}{2}t^2} = \sum_{n=0}^\infty H_n(x)\frac{t^n}{n!}.$$

We can readily obtain that the above basis is equivalent to the Hermite expansion $H_n(\xi_k)$'s.

2.3. **Wiener chaos methods for linear parabolic equations.** Consider the linear SPDE in Ito's form over $(t,x) \in (0,T] \times \mathcal{D}$,

$$\begin{aligned} du(t,x) &= [\mathcal{L}u(t,x) + f(x)]\,dt + \sum_{k=1}^q [\mathcal{M}_k u(t,x) + g_k(x)]\,dw_k(t), \\ u(0,x) &= u_0(x), \quad x \in \mathcal{D}, \end{aligned}$$

where $\mathcal{D} \subsetneq \mathbb{R}^d$ and $\{w_k(t), 1 \le k \le q\}$ q-dimensional Wiener process and certain boundary conditions for wellposedness are imposed.

Step 1. We need to write the Ito formulation (with Wick product): $t \in (0,T]$

$$\frac{d}{dt}u(t,x) = \mathcal{L}u(t,x) + f(x) + \sum_{k=1}^q [\mathcal{M}_k u(t,x) + g_k(x)] \diamond \sum_{i=1}^{\mathsf{n}} m_i(t)\xi_{i,k}.$$

Step 2. We apply the Hermite spectral approximation in random space: $u = \sum_{\alpha\in\mathcal{J}_{\mathsf{N},\mathsf{n}}} \varphi_\alpha \xi_\alpha$.

Together with the Ito-Wick product, we have

$$\begin{aligned}\frac{\partial\varphi_\alpha(t,x;\phi)}{\partial t} &= \mathcal{L}\varphi_\alpha(t,x;\phi) + f(x)\mathbf{1}_{\{|\alpha|=0\}} \\ &\quad + \sum_{k=1}^{q}\sum_{l=1}^{\mathsf{n}} \sqrt{\alpha_{k,l}}\, m_l(s)\left[\mathcal{M}_k\varphi_{\alpha^-(k,l)}(s,x;\phi) + g_k(x)\mathbf{1}_{\{|\alpha|=1\}}\right], \quad s\in(0,T], \\ \varphi_\alpha(0,x) &= \phi(x)\mathbf{1}_{\{|\alpha|=0\}},\end{aligned}$$

where $\alpha^-(k,l)$ is the multi-index with components

$$\left(\alpha^-(k,l)\right)_{i,j} = \begin{cases} \max(0,\alpha_{i,j}-1), & \text{if } i=k \text{ and } j=l, \\ \alpha_{i,j}, & \text{otherwise.}\end{cases}$$

Observe that this linear system is actually lower-triangular.

Step 3. Now apply your favorite PDE solver(s) for the propagator.

Step 4. Compute the statistics of the numerical solution using the truncated Wiener chaos expansion. For example, the mean is ψ_0 and the second-order moment is $\mathbb{E}[u^2] = \sum_{\alpha\in\mathcal{J}_{\mathsf{N},\mathsf{n}}} \varphi_\alpha^2$. The higher-order moments can be also computed, e.g.,

$$\mathbb{E}[u^3(t,x)] = \sum_{\gamma\in\mathcal{J}}\sum_{(0)\le\beta\le\alpha} \Phi(\alpha,\beta,\gamma)\psi_{\alpha-\beta+\gamma}\psi_{\beta+\gamma}\psi_\alpha,$$

where

$$\Phi(\alpha,\beta,\gamma) = \left[\binom{\alpha}{\beta}\binom{\beta+\gamma}{\gamma}\binom{\alpha-\beta+\gamma}{\gamma}\right]^{1/2}. \tag{2.14}$$

The key for this calculation is to observe that

$$\xi_\alpha\xi_\beta = \sum_{\gamma\le\alpha\wedge\beta} B(\alpha,\beta,\gamma)\xi_{\alpha+\beta-2\gamma}, \tag{2.15}$$

where $\Phi(\alpha,\beta,\gamma) = B(\beta+\gamma,\alpha-\beta+\gamma,\gamma)$ and

$$B(\alpha,\beta,\gamma) = \left(\binom{\alpha}{\gamma}\binom{\beta}{\gamma}\binom{\alpha+\beta-2\gamma}{k-\gamma}\right)^{\frac{1}{2}}. \tag{2.16}$$

Fourth-order moments can be also derived.

Remark 2.3. *The WCE can also be applied to nonlinear equations, see, e.g., Chapter 11 of* [Zhang and Karniadakis, 2017]*. However, the direct application of WCE only works for a very short-time integration. To achieve a longer time integration, please refer to Chapter 6 of* [Zhang and Karniadakis, 2017].

3. What could go wrong: regularity matters

As a special class of parabolic SPDEs, stochastic Burgers and Navier-Stokes equations require more attention for their strong interactions between the strong nonlinearity and the noises. Similar to the linear heat equation with additive noise, the convergence for the time-discretization of one-dimensional Burgers equations is no more than 1/4.

Because of the strong nonlinearity, the discretization in space and in time may cause some effects, such as "a spatial version of the Ito-Stratonovich correction" [Hairer and Maas, 2012, Hairer and Voss, 2011]. Hairer et al considered finite difference schemes for the Burgers equation with additive space-time noise in [Hairer and Voss, 2011]:

$$\partial_t u = \nu\partial_x^2 u + (\nabla G(u))\partial_x u + \sigma\dot{W}^Q, \quad x\in[0,2\pi]. \tag{3.1}$$

If we only consider the discretization of the first-order differential operator, e.g.,

$$\partial_t u^\varepsilon = \nu \partial_x^2 u^\varepsilon + (\nabla G(u^\varepsilon))\partial_x u^\varepsilon + \sigma \dot{W}^Q, \quad \partial_x u^\varepsilon =: \frac{u(x+a\varepsilon) - u(x-b\varepsilon)}{(a+b)\varepsilon}, \; a, b > 0, \tag{3.2}$$

then it can be proved that this equation converges to (see [Hairer and Maas, 2012])

$$\partial_t v = \nu \partial_x^2 v + (\nabla G(v))\partial_x v - \frac{\sigma^2}{4\nu}\frac{a-b}{a+b}\Delta G(v) + \sigma \dot{W}^Q, \quad x \in [0, 2\pi], \tag{3.3}$$

if $\dot{W}^Q$ is space-time white noise; and no correction term appears if $\dot{W}^Q$ is more regular than space-time white noise, e.g. white in time but correlated in space. Effects of some other standard discretizations in space, e.g. Galerkin methods, and fully discretizations were also discussed in [Hairer and Maas, 2012].

References

[Hairer and Maas, 2012] Hairer, M. and Maas, J. (2012). A spatial version of the Itô-Stratonovich correction. *Ann. Probab.*, 40(4):1675–1714.

[Hairer and Voss, 2011] Hairer, M. and Voss, J. (2011). Approximations to the stochastic Burgers equation. *J. Nonlinear Sci.*, 21(6):897–920.

[Milstein and Tretyakov, 2009] Milstein, G. N. and Tretyakov, M. V. (2009). Solving parabolic stochastic partial differential equations via averaging over characteristics. *Math. Comp.*, 78(268):2075–2106.

[Rozovsky and Lototsky, 2018] Rozovsky, B. L. and Lototsky, S. V. (2018). *Stochastic evolution systems*, volume 89 of *Probability Theory and Stochastic Modelling*. Springer, Cham. Linear theory and applications to non-linear filtering, Second edition of [MR1135324].

[Zhang and Karniadakis, 2017] Zhang, Z. and Karniadakis, G. E. (2017). *Numerical methods for stochastic partial differential equations with white noise*, volume 196 of *Applied Mathematical Sciences*. Springer, Cham.