Tensor methods for deterministic and stochastic PDEs

We have seen how to propagate uncertainty in PDE models involving random initial conditions, random parameters, random boundary conditions or random forcing terms. Specifically, we disussed PDF methods (Hopf equation and Lundgren-Monin-Novikov hierarchies), polynomial chaos methods (gPC, ME-gPC), and sampling methods (MC, qMC, PCM, ME-PCM, and sparse grids). In this lecture note we discuss another method that relies upon orthogonal tensor expansions to compute the solution of stochastic PDEs. The same theory can be used to compute the numerical solution of high-dimensional PDEs such as the Liouville equation or the Fokker-Plank equation.

Dynamically orthogonal (DO) tensor methods for stochastic PDEs

The dynamically orthogonal field equation method for SPDEs was pioneered by Sapsis and Lermusiaux in [14], and it is essentially a tensor method for linear or nonlinear PDEs in a separable Hilbert space [7, 6]. To describe DO, suppose we are interested in computing the solution to an initial/boundary value problem for a stochastic PDE of the form

$$\begin{cases} \dfrac{\partial u(\boldsymbol{x},t;\omega)}{\partial t} = G_\omega(u(\boldsymbol{x},t;\omega)), \\ u(\boldsymbol{x},0;\omega) = u_0(\boldsymbol{x};\omega), \end{cases} \tag{1}$$

where $\boldsymbol{x} \in V \subseteq \mathbb{R}^d$ (V is the spatial domain $d \geq 1$), and G_ω is a *random nonlinear operator* which may take into account random forcing terms, random parameters or random boundary conditions. A simple example of $G_\omega(u(\boldsymbol{x},t;\omega))$ could be[1]

$$G_\omega\left(u(\boldsymbol{x},t;\omega) = \nabla \cdot \left[\kappa(\boldsymbol{x};\omega)\nabla u(\boldsymbol{x},t;\omega)\right], \qquad \kappa(\boldsymbol{x};\omega) > 0, \right. \tag{2}$$

in $\mathbb{R}^d$. We look for a representation of the solution to (1) of the form

$$u(\boldsymbol{x},t;\omega) = \mathbb{E}\{u(\boldsymbol{x},t;\omega)\} + \sum_{k=1}^{\infty} \widehat{u}_k(\boldsymbol{x},t) Y_k(t;\omega), \tag{3}$$

where $\{\widehat{u}_1(\boldsymbol{x},t), \widehat{u}_2(\boldsymbol{x},t), \cdots\}$ are deterministic spatio-temporal modes, while $\{Y_1(t;\omega), Y_2(t;\omega), \ldots\}$ are random temporal modes. Note the time redundancy in both the space-time modes $\widehat{u}_k(\boldsymbol{x},t)$ and the random modes $Y_k(t;\omega)$. The theoretical justification of the series expansion (3) relies on a tensor product representation of the Hilbert space $L^2(V \times T \times \Omega)$ (T is the temporal domain and Ω is the sample space) as

$$L^2(V \times T \times \Omega) = L^2(V \times T) \otimes L^2(T \times \Omega). \tag{4}$$

The expansion (3) includes time-dependent gPC [9] as a sub-case.

Properties of the modes $\widehat{u}_k(\boldsymbol{x},t)$ and $Y_k(t;\omega)$. The random temporal modes $Y_k(t;\omega)$ are clearly zero mean. In fact, by applying the expectation operator to (3) we obtain

$$\sum_{k=1}^{\infty} \widehat{u}_k(\boldsymbol{x},t)\mathbb{E}\{Y_k(t;\omega)\} = 0 \quad \Rightarrow \quad \mathbb{E}\{Y_k(t;\omega)\} = 0. \tag{5}$$

[1] The PDE (1)-(2) describes heat conduction in a heterogeneous medium with random thermal conductivity $\kappa(\boldsymbol{x};\omega)$.

We also assume that the space-time modes $\widehat{u}_k(\boldsymbol{x},t)$ satisfy the *gauge*[2] conditions

$$\left\langle \widehat{u}_k(\boldsymbol{x},t), \frac{\partial \widehat{u}_j(\boldsymbol{x},t)}{\partial t} \right\rangle_{L^2(V)} = \int_V \widehat{u}_k(\boldsymbol{x},t) \frac{\partial \widehat{u}_j(\boldsymbol{x},t)}{\partial t} d\boldsymbol{x} = 0 \qquad \text{for all } t \geq 0 \qquad \text{and all } j,k \geq 1. \tag{6}$$

These conditions are called *dynamically orthogonal (DO) conditions*. The reason is that if the set of modes $\{\widehat{u}_k(\boldsymbol{x},t)\}$ is initially orthonormal, i.e.,

$$\langle \widehat{u}_k(\boldsymbol{x},0), \widehat{u}_j(\boldsymbol{x},0) \rangle_{L^2(V)} = \delta_{kj}. \tag{7}$$

then it stays orthonormal in time. In fact, for all $t \geq 0$ we have

$$\frac{\partial}{\partial t} \langle \widehat{u}_k(\boldsymbol{x},t), \widehat{u}_j(\boldsymbol{x},t) \rangle_{L^2(V)} = \left\langle \frac{\partial \widehat{u}_k(\boldsymbol{x},t)}{\partial t}, \widehat{u}_j(\boldsymbol{x},t) \right\rangle_{L^2(V)} + \left\langle \widehat{u}_k(\boldsymbol{x},t), \frac{\partial \widehat{u}_j(\boldsymbol{x},t)}{\partial t} \right\rangle_{L^2(V)} = 0 \quad \text{for all } i,j \geq 1. \tag{8}$$

This implies that

$$\langle \widehat{u}_k(\boldsymbol{x},t), \widehat{u}_j(\boldsymbol{x},t) \rangle_{L^2(V)} = \langle \widehat{u}_k(\boldsymbol{x},0), \widehat{u}_j(\boldsymbol{x},0) \rangle_{L^2(V)} = \delta_{kj}, \tag{9}$$

i.e., space time modes $\widehat{u}_k(\boldsymbol{x},t)$ that are orthogonal at $t = 0$ remain orthogonal at later times. For this reason we shall call $\widehat{u}_k(\boldsymbol{x},t)$ *dynamically orthogonal* modes.

DO propagator. At this point we have all elements to derive a coupled system of equations for the DO modes $\widehat{u}_j(\boldsymbol{x},t)$, the stochastic modes $Y_t(t;\omega)$ and the mean field

$$\overline{u}(\boldsymbol{x},t) = \mathbb{E}\{u(\boldsymbol{x},t;\omega\} \tag{10}$$

appearing in (3). To this end, we first substitute a truncated expansion of the form (3), i.e.,

$$u_M(\boldsymbol{x},t;\omega) = \overline{u}(\boldsymbol{x},t) + \sum_{k=1}^{M} \widehat{u}_k(\boldsymbol{x},t) Y_k(t;\omega), \tag{11}$$

into the SPDE (1) to obtain

$$\frac{\partial \overline{u}(\boldsymbol{x},t)}{\partial t} + \sum_{k=1}^{M} \left(\frac{\partial \widehat{u}_k(\boldsymbol{x},t)}{\partial t} Y_k(t;\omega) + \widehat{u}_k(\boldsymbol{x},t) \frac{dY_k(t;\omega)}{dt} \right) = G_\omega\left(u_M(\boldsymbol{x},t;\omega)\right) + R_M(\boldsymbol{x},t;\omega). \tag{12}$$

Then we impose that the residual $R_M(\boldsymbol{x},t;\omega)$ is orthogonal to

$$S_M = \text{span}\left\{\widehat{u}_1(\boldsymbol{x},t), \ldots, \widehat{u}_M(\boldsymbol{x},t)\right\} \quad \text{and} \quad Z_M = \text{span}\left\{Y_1(t;\omega), \ldots, Y_M(t;\omega)\right\} \tag{13}$$

relative to the inner products $\langle\cdot\rangle_{L^2(V)}$ (see Eq. (6)) and $\mathbb{E}\{\cdot\}$. This gives the $2M+1$ conditions

$$0 = \mathbb{E}\left\{R_M(\boldsymbol{x},t;\omega)\right\}, \tag{14}$$

$$0 = \mathbb{E}\left\{R_M(\boldsymbol{x},t;\omega) Y_k(t;\omega)\right\} \qquad k = 1,\ldots,M, \tag{15}$$

$$0 = \langle R_M(\boldsymbol{x},t;\omega) \widehat{u}_k(\boldsymbol{x},t) \rangle_{L^2(V)} \qquad k = 1,\ldots,M \tag{16}$$

which are sufficient to identify a set of equation for the mean field $\boldsymbol{u}(\boldsymbol{x},t)$, the DO modes $\{\widehat{u}_k(\boldsymbol{x},t)\}$, and the stochastic modes $\{Y_k(t;\omega)\}$. By taking the expectation of (12) and taking into account (14) we obtain

$$\frac{\partial \overline{u}}{\partial t} = \mathbb{E}\left\{G_\omega\left(u_M\right)\right\} \quad \text{(evolution equation for the mean field)}. \tag{17}$$

[2]In physics, choosing a gauge denotes a mathematical procedure for coping with redundant degrees of freedom in field variables. In the case of the series expansion (3), t is the redundant degree of freedom. We also emphasize that the inner product (6) can be generalized to include, e.g., a weight function $\mu(\boldsymbol{x})$ (weighted $L^2_\mu(V)$ space), or spatial derivatives of $\widehat{u}_k(\boldsymbol{x},t)$ (Sobolev space $H^s(V)$).

Next, we project (12) onto $\widehat{u}_p(\boldsymbol{x},t)$ and take (16) into account to obtain

$$\left\langle \frac{\partial \overline{u}}{\partial t}, \widehat{u}_p \right\rangle_{L^2(V)} + \sum_{k=1}^{M} \underbrace{\left\langle \frac{\partial \widehat{u}_k}{\partial t}, \widehat{u}_p \right\rangle_{L^2(V)}}_{=0} Y_k + \sum_{k=1}^{M} \underbrace{\langle \widehat{u}_k \widehat{u}_p \rangle_{L^2(V)}}_{=\delta_{kp}} \frac{dY_k}{dt} = \langle G_\omega\left(u_M\right) \widehat{u}_p \rangle_{L^2(V)}\,, \tag{18}$$

where we assumed that the DO modes $\{\widehat{u}_k(\boldsymbol{x},t)\}$ are orthonormal at $t=0$ and therefore at all t (see Eq. (9)). Equation (18) can be written as

$$\frac{dY_p}{dt} = \langle [G_\omega\left(u_M\right) - \mathbb{E}\{G_\omega\left(u_M\right)\}] \widehat{u}_p \rangle_{L^2(V)}\,. \tag{19}$$

Finally we project (12) onto $Y_p(t;\omega)$ and take (16) into account to obtain

$$\underbrace{\mathbb{E}\left\{ \frac{\partial \overline{u}}{\partial t} Y_p \right\}}_{=0} + \sum_{k=1}^{M} \frac{\partial \widehat{u}_k}{\partial t} \underbrace{\mathbb{E}\left\{ Y_k Y_p \right\}}_{\Sigma_{kp}(t)} + \sum_{k=1}^{M} \widehat{u}_k \mathbb{E}\left\{ \frac{dY_k}{dt} Y_p \right\} = \mathbb{E}\left\{ G_\omega\left(u_M\right) Y_p \right\}. \tag{20}$$

Note that

$$\Sigma_{kp}(t) = \mathbb{E}\left\{ Y_k(t;\omega) Y_p(t;\omega) \right\} \tag{21}$$

is the *covariance function* of the random process $Y_k(t;\omega)$ and $Y_p(t;\omega)$. By using (19) we can write the last terms at the right hand side of (20) as[3]

$$\mathbb{E}\left\{ \frac{dY_k}{dt} Y_p \right\} = \mathbb{E}\left\{ \langle G_\omega(u_M) \widehat{u}_k \rangle_{L^2(V)} Y_p \right\}. \tag{23}$$

A substitution of (23) into (20) yields

$$\sum_{k=1}^{M} \frac{\partial \widehat{u}_k}{\partial t} \Sigma_{kp}(t) = \mathbb{E}\left\{ G_\omega\left(u_M\right) Y_p \right\} - \sum_{k=1}^{M} \langle \mathbb{E}\left\{ G_\omega\left(u_M\right) Y_p \right\} \widehat{u}_k \rangle_{L^2(V)} \widehat{u}_k. \tag{24}$$

In summary, the *DO propagator* can be written as [14, 5] (for $p = 1, \ldots, M$)

$$\begin{cases} \dfrac{\partial \overline{u}}{\partial t} = \mathbb{E}\left\{ G_\omega\left(u_M\right) \right\}, \\ \dfrac{dY_p}{dt} = \langle [G_\omega\left(u_M\right) - \mathbb{E}\{G_\omega\left(u_M\right)\}] \widehat{u}_p \rangle_{L^2(V)}\,, \\ \displaystyle\sum_{k=1}^{M} \frac{\partial \widehat{u}_k}{\partial t} \Sigma_{kp}(t) = \mathbb{E}\left\{ G_\omega\left(u_M\right) Y_p \right\} - \sum_{k=1}^{M} \langle \mathbb{E}\left\{ G_\omega\left(u_M\right) Y_p \right\} \widehat{u}_k \rangle_{L^2(V)} \widehat{u}_k. \end{cases} \tag{25}$$

The initial and boundary conditions for this PDE system are obtained by projection (see [14]). Clearly, the evolution equations for the DO modes $\widehat{u}_k$ in (25) have some issues if the covariance matrix Σ_{kp} of the stochastic modes is singular. This happens, for example when a random mode Y_k has zero energy, e.g., when we add a mode during integration to increase accuracy. In this case, the system (25) becomes algebraic-differential (covariance matrix singular). This requires special numerical techniques for temporal integration. One can overcome this problem by considering pseudo-inverse matrix operations [1]. More rigorously, it can be shown that is possible to rewrite the system (25) in fully equivalent form that does not require covariance matrix inversion, and solve such a system using *operator splitting* (see, e.g., [6]).

[3] Note that

$$\mathbb{E}\left\{ \mathbb{E}\left\{ G_\omega(u_M) \right\} Y_p \right\} = \mathbb{E}\left\{ G_\omega(u_M) \right\} \mathbb{E}\left\{ Y_p \right\} = 0. \tag{22}$$

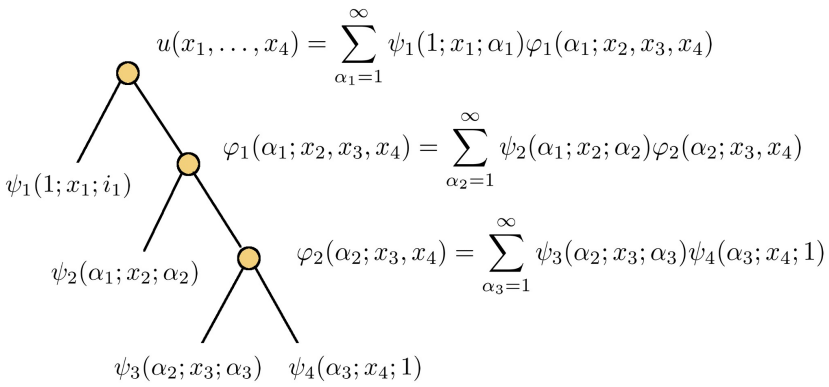

Figure 1: Construction of functional tensor train (FTT). Shown is the sequence of hierarchical Schmidt decomposition for a function in four variables.

An advantage of (25) over, e.g., polynomial chaos is that the stochastic modes are evolving with time in a way that depends on the PDE. Moreover, it can be shown that the DO equations (25) satisfy an optimality (variational) principle similar to the one satisfied by the Karhunen-Loève expansion (see [7]), which implies that we can obtain accurate stochastic solutions of (1) using a series expansion (11) with a relatively small number of modes M.

In a parallel research effort, T. Hou and collaborators developed an alternative version of DO on bi-orthogonal (BO) expansions [3, 4]. Bi-orthogonality essentially represents a different *gauge* condition, which yields a propagator, i.e., a coupled system of equations for the modes $\widehat{u}_k$ and Y_k that differs from (25)(see [3, 4] for details). The correspondence between DO and BO was investigated in [5, 7].

Dynamically orthogonal tensor methods for high-dimensional deterministic PDE

In this section we generalize the series expansion (3) to compute the numerical solution of a high-dimensional deterministic PDE of the form

$$\frac{\partial u(\boldsymbol{x},t)}{\partial t}=G(u(\boldsymbol{x},t)),\qquad u(\boldsymbol{x},0)=u_0(\boldsymbol{x}), \tag{26}$$

where $u:V\times[0,T]\to\mathbb{R}$ is a (time-dependent) scalar field in d variables defined on the domain $V\subseteq\mathbb{R}^d$ and G is a nonlinear operator which may depend on the spatial variables, and may incorporate boundary conditions.

The PDE (26) may be a Liouville equation, a Fokker-Planck equation, or an approximation of the Hopf characteristic functional equation we have seen in Chapter 2.

Functional tensor train (FTT). Let $V\subseteq\mathbb{R}^d$ be a Cartesian product of d real intervals $V_i=[a_i,b_i]$

$$V=\bigtimes_{i=1}^{d}V_i, \tag{27}$$

μ a finite product measure on V

$$\mu(\boldsymbol{x})=\prod_{i=1}^{d}\mu_i(x_i), \tag{28}$$

and

$$H=L^2_\mu(V) \tag{29}$$

the standard weighted Hilbert space[4] of square–integrable functions on V. It was shown in [12, 2, 8] that any function $u(\boldsymbol{x}) \in H$ can be represented as

$$u(\boldsymbol{x}) = \sum_{\alpha_1,\ldots,\alpha_{d-1}=1}^{\infty} \psi_1(1;x_1;\alpha_1)\psi_2(\alpha_1;x_2;\alpha_2)\cdots\psi_d(\alpha_{d-1};x_d;1), \tag{30}$$

where $\psi_i(\alpha_{i-1};x_i;\alpha_i)$ are matrices of functions depending only on the variable x_i. Such functions are computed by solving a hierarchical sequence of eigenvalue problems that is similar to the Karhunen-Loève eigenvalue problem.

Computation of FTT. In Figure 1 we show the sequence of hierarchical (Schmidt) decompositions to compute the functional tensor train expansion for a four-dimensional function. The first step is to solve the eigenvalue problem

$$\lambda_1\psi_1(1;x_1;\alpha_1) = \int_{V_1} K(x_1,x_1')\psi_1(1;x_1';\alpha_1)dx_1', \tag{31}$$

where

$$K_1(x_1,x_1') = \int_{V_2\times V_3\times V_4} u(x_1,x_2,x_3,x_4)u(x_1',x_2,x_3,x_4)dx_2dx_3dx_4. \tag{32}$$

The (not-normalized) modes $\varphi_1(\alpha_1;x_2,x_3,x_4)$ are obtained by projection of u onto the orthonormal modes ψ_1 as

$$\varphi_1(\alpha_1;x_2,x_3,x_4) = \int_{V_1} u(x_1,x_2,x_3,x_4)\psi_1(1;x_1;\alpha_1)dx_1. \tag{33}$$

At this point we perform another Schmidt decomposition by solving the eigenvalue problem

$$\lambda_2\psi_2(\alpha_1;x_2;\alpha_2) = \int_{V_2} K_2(x_2,x_2';\alpha_1)\psi_2(\alpha_1;x_2';\alpha_2)dx_2', \tag{34}$$

where

$$K_2(x_2,x_2';\alpha_1) = \int_{V_3\times V_4} \varphi_1(\alpha_1;x_2,x_3,x_4)\varphi_1(\alpha_1;x_2',x_3,x_4)dx_3dx_4. \tag{35}$$

Note that the kernel K_2 is defined by the orthogonal modes φ_1 we obtained from the previous decomposition. We project $\varphi_1(\alpha_1;x_2,x_3,x_4)$ onto the orthonormal modes $\psi_2(\alpha_1;x_2';\alpha_2)$ to obtain

$$\varphi_2(\alpha_2;x_3,x_4) = \sum_{\alpha_1=1}^{\infty}\int_{V_2} \varphi_1(\alpha_1;x_2,x_3,x_4)\psi_2(\alpha_1;x_2;\alpha_2)dx_2. \tag{36}$$

Lastly we perform a decomposition the $\varphi_2(\alpha_2;x_3,x_4)$, which yields the modes $\psi_3(\alpha_2;x_3;\alpha_3)$ and $\psi_4(\alpha_3;x_4;1)$ (see Figure 1). The final expansion corresponds to the following sequence of function space decompositions

$$\begin{aligned} H(V_1\times V_2\times V_3\times V_4) =& [H(V_1)\otimes H(V_2\times V_3\times V_4)] \\ & [H(V_1)\otimes [H(V_2)\otimes H(V_3\times V_4)]] \\ & [H(V_1)\otimes [H(V_2)\otimes [H(V_3)\otimes H(V_4)]]], \end{aligned} \tag{37}$$

where the notation $[H(V_1)\otimes H(V_2\times V_3\times V_4)]$ emphasizes the fact that that we diagonalized the expansion involving the function spaces within the bracket.

In a finite-dimensional setting, such decomposition are essentially generated by a hierarchical sequence of singular value decompositions corresponding to various flattening of a multi-dimensional array (see Figure

[4]Note that the Hilbert space H in equation (29) can be equivalently chosen to be a Sobolev space $W^{2,p}$ (see [7] for details).

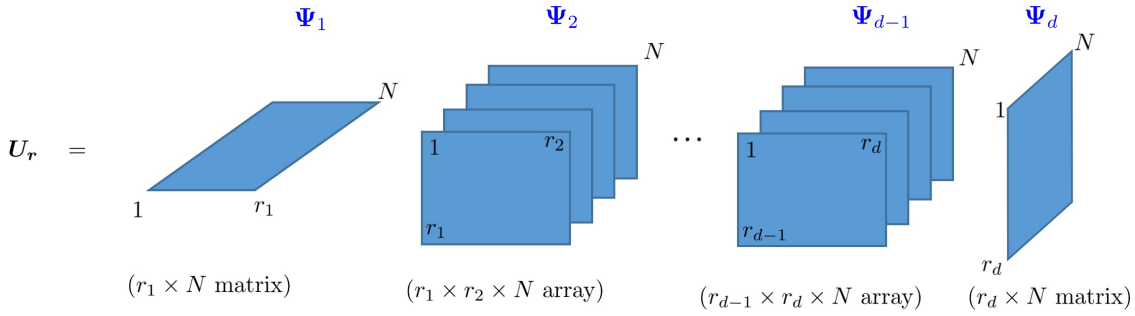

Figure 2: Construction of functional tensor train (FTT). Shown is the sequence of hierarchical Schmidt decomposition for a function in four variables.

2). By truncating the expansion (30) so that the terms corresponding to the largest eigenvalues in (31), (34), etc., are retained yields

$$u_{\boldsymbol{r}}(\boldsymbol{x}) = \sum_{\alpha_1,\ldots,\alpha_{d-1}=1}^{\boldsymbol{r}} \psi_1(1;x_1;\alpha_1)\psi_2(\alpha_1;x_2;\alpha_2)\cdots\psi_d(\alpha_{d-1};x_d;1), \tag{38}$$

where $\boldsymbol{r} = (1, r_1, \ldots, r_{d-1}, 1)$ is the FTT-rank.

It was shown by Bigoni *et al.* [2] that the truncated FTT expansion expansion (38) converges optimally (in $\mathbf{r}$) with respect to the $L^2_\mu(V)$ norm. More precisely, for any given function $u \in L^2_\mu(V)$ the FTT approximant (38) minimizes the residual $\|u - u_{\boldsymbol{r}}\|_{L^2_\mu(V)}$ relative to independent variations of the functions $\psi_i(\alpha_{i-1}; x_i; \alpha_i)$ on a tensor manifold with constant rank $\boldsymbol{r}$. It is convenient to write (38) in a more compact form as

$$u_{\boldsymbol{r}}(\boldsymbol{x}) = \boldsymbol{\Psi}_1(x_1)\boldsymbol{\Psi}_2(x_2)\cdots\boldsymbol{\Psi}_d(x_d), \tag{39}$$

where $\boldsymbol{\Psi}_i(x_i)$ is a $r_{i-1}\times r_i$ matrix with entries $[\boldsymbol{\Psi}_i(x_i)]_{jk} = \psi_i(j; x_i; k)$. The matrix-valued functions $\boldsymbol{\Psi}_i(x_i)$ are known as FTT *tensor cores*. To simplify notation even more we can suppress explicit tensor core dependence on the spatial variable x_i, allowing us to simply write $\boldsymbol{\Psi}_i = \boldsymbol{\Psi}_i(x_i)$ as the spatial dependence is indicated by the tensor core subscript. If we discretize the domain V in terms of a grid with N points in each variable then we can represent (39) as a product of 2D and 3D matrices (see Figure 2).

FTT tensor manifold. It was shown in [15, 8] that the set of truncated tensors (38) (with invertible covariance matrices of each tensor modes) belongs to a *smooth manifold*[5] $\mathcal{M}_{\boldsymbol{r}}$, i.e., a manifold in which we can define a tangent space $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ at a point $u_{\boldsymbol{r}} \in \mathcal{M}_{\boldsymbol{r}}$. Specifically, let us denote by $H^{(i)}_{r_{i-1}\times r_i}$ the set of all tensor cores $\boldsymbol{\Psi}_i \in M_{r_{i-1}\times r_i}(L^2_{\mu_i}(V_i))$ with the property that the autocovariance matrices $\langle \boldsymbol{\Psi}_i^{\mathrm{T}}\boldsymbol{\Psi}_i\rangle_i \in M_{r_i\times r_i}(\mathbb{R})$ and $\langle \boldsymbol{\Psi}_i\boldsymbol{\Psi}_i^{\mathrm{T}}\rangle_i \in M_{r_{i-1}\times r_{i-1}}(\mathbb{R})$ are invertible for $i = 1, \ldots, d$. The set

$$\mathcal{M}_{\boldsymbol{r}} = \{u_{\boldsymbol{r}} \in L^2_\mu(V): \quad u_{\boldsymbol{r}} = \boldsymbol{\Psi}_1\boldsymbol{\Psi}_2\cdots\boldsymbol{\Psi}_d, \quad \boldsymbol{\Psi}_i \in H^{(i)}_{r_{i-1}\times r_i}, \quad \forall i = 1, 2, \ldots, d\}, \tag{40}$$

consisting of fixed-rank FTT tensors, is a smooth sub-manifold of $L^2_\mu(V)$. We represent elements in the tangent space, $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$, of $\mathcal{M}_{\boldsymbol{r}}$ at the point $u_{\boldsymbol{r}} \in \mathcal{M}_{\boldsymbol{r}}$ as equivalence classes of velocities of continuously differentiable curves on $\mathcal{M}_{\boldsymbol{r}}$ passing through $u_{\boldsymbol{r}}$

$$T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}} = \left\{\gamma'(s)|_{s=0}: \quad \gamma \in \mathcal{C}^1\left((-\delta,\delta), \mathcal{M}_{\boldsymbol{r}}\right), \quad \gamma(0) = u_{\boldsymbol{r}}\right\}. \tag{41}$$

[5] A manifold is a generalization and abstraction of the notion of a curved surface. In particular, the manifold of the FTT tensors with fixed rank is a topological function space that admits a tangent space at each point, an inner product defined on the tangent space, etc.

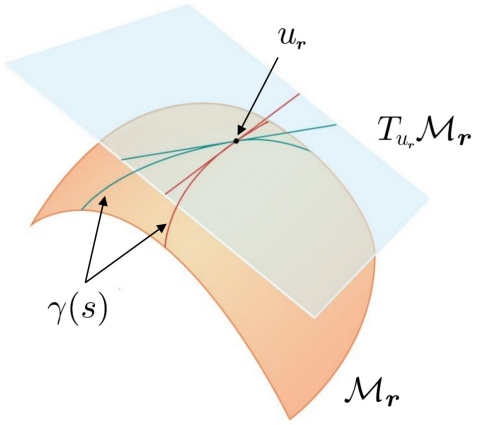

Figure 3: Sketch of the tensor manifold $\mathcal{M}_{\boldsymbol{r}}$ and the tangent space $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ at $u_{\boldsymbol{r}} \in \mathcal{M}_{\boldsymbol{r}}$. The tangent space is defined as equivalence classes of velocities of continuously differentiable curves $\gamma(s)$ on $\mathcal{M}_{\boldsymbol{r}}$ passing through $u_{\boldsymbol{r}}$.

A sketch of $\mathcal{M}_{\boldsymbol{r}}$ and $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ is provided in Figure 3. Since $L^2_\mu(V)$ is an inner product space, for each $u \in L^2_\mu(V)$ the tangent space $T_u L^2_\mu(V)$ is canonically isomorphic to $L^2_\mu(V)$. Moreover, for each $u_{\boldsymbol{r}} \in \mathcal{M}_{\boldsymbol{r}}$ the normal space to $\mathcal{M}_{\boldsymbol{r}}$ at the point $u_{\boldsymbol{r}}$, denoted by $N_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$, consists of all vectors in $L^2_\mu(V)$ that are orthogonal to $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ with respect to the inner product in $L^2_\mu(V)$

$$N_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}} = \{w \in L^2_\mu(V) : \langle w, v\rangle_{L^2_\mu(V)} = 0, \quad \forall v \in T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}\}. \tag{42}$$

Since the tangent space $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ is closed, for each point $u_{\boldsymbol{r}} \in \mathcal{M}_{\boldsymbol{r}}$ the space $L^2_\mu(V)$ admits a decomposition into tangential and normal components

$$L^2_\mu(V) = T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}} \oplus N_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}. \tag{43}$$

We represent elements of the tangent space $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ as equivalence classes of velocities of curves passing through the point $u_{\boldsymbol{r}}$

$$T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}} = \left\{y'(s)|_{s=0} : \quad y \in \mathcal{C}^1\left((-\delta,\delta), \mathcal{M}_{\boldsymbol{r}}\right), \quad y(0) = u_{\boldsymbol{r}}\right\}. \tag{44}$$

Here $\mathcal{C}^1\left((-\delta,\delta), \mathcal{M}_{\boldsymbol{r}}\right)$ is the space of continuously differentiable functions from the interval $(-\delta,\delta)$ to the space of constant rank FTT tensors $\mathcal{M}_{\boldsymbol{r}}$.

Next, we can now define a projection onto the tangent space of $\mathcal{M}_{\boldsymbol{r}}$ at $u_{\boldsymbol{r}}$ by

$$\begin{aligned} P_{u_{\boldsymbol{r}}} &: L^2_\mu(V) \to T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}} \\ P_{u_{\boldsymbol{r}}} v &= \operatorname*{argmin}_{v_{\boldsymbol{r}} \in T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}} \|v - v_{\boldsymbol{r}}\|_{L^2_\mu(V)}. \end{aligned} \tag{45}$$

For fixed $u_{\boldsymbol{r}}$, the map $P_{u_{\boldsymbol{r}}}$ is linear and bounded. Each $v \in L^2_\mu(V)$ admits a unique representation as $v = v_t + v_n$ where $v_t \in T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ and $v_n \in N_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ (see equation (43)). From this representation it is clear that $P_{u_{\boldsymbol{r}}}$ is an orthogonal projection onto the tangent space $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$.

An arbitrary element of the tangent space $T_{u_{\boldsymbol{r}}}\mathcal{M}_{\boldsymbol{r}}$ can be expressed as

$$\dot{u}_{\boldsymbol{r}} = \dot{\boldsymbol{\Psi}}_1 \boldsymbol{\Psi}_{\geq 2} + \cdots + \boldsymbol{\Psi}_{\leq i-1} \dot{\boldsymbol{\Psi}}_i \boldsymbol{\Psi}_{\geq i+1} + \cdots + \boldsymbol{\Psi}_{\leq d-1} \dot{\boldsymbol{\Psi}}_d, \tag{46}$$

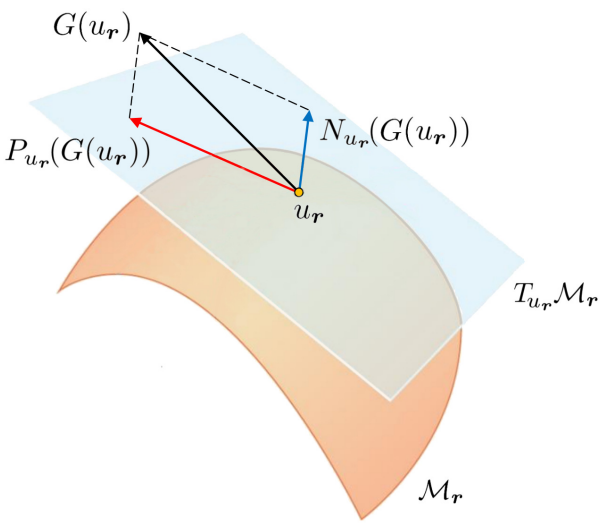

Figure 4: Tangent and normal components of $G\left(u_{\boldsymbol{r}}\right) = \partial u_{\boldsymbol{r}}/\partial t$ at $u_{\boldsymbol{r}}$. The tensor rank of the solution is increased at time t_i if the norm of the normal component $N_{u_{\boldsymbol{r}}}(G(u_{\boldsymbol{r}}))$ is larger than a specified threshold ϵ_{inc}.

where $\dot{u}_{\boldsymbol{r}} = \partial u_{\boldsymbol{r}}/\partial t$ and $\dot{\boldsymbol{\Psi}}_i = \partial \boldsymbol{\Psi}_i/\partial t$.

Dynamic tensor approximation of high-dimensional nonlinear PDEs. With the machinery on FTT tensors available, we can now approximate the solution of (26) on the tensor manifold $\mathbb{M}_{\boldsymbol{r}}$. To this end, suppose that the initial condition $u_0(\boldsymbol{x})$ is on the manifold $\mathcal{M}_{\boldsymbol{r}}$. Clearly, the solution to the initial/boundary value problem (see Figure 4)

$$\begin{cases} \dfrac{\partial u_{\boldsymbol{r}}}{\partial t} = P_{u_{\boldsymbol{r}}} G(u_{\boldsymbol{r}}), \\ u(\boldsymbol{x},0) = u_0(\boldsymbol{x}), \end{cases} \tag{47}$$

remains on the manifold $\mathcal{M}_{\boldsymbol{r}}$ for all $t \geq 0$. Here G is the nonlinear operator on the right hand side of equation (1). The solution to (47) is known as a dynamic approximation to the solution of (1). To compute the tangent space projection of the PDE (48) we solve the *convex optimization problem*

$$\min_{v(\boldsymbol{x},t)\in T_{u(\boldsymbol{x},t)}\mathcal{M}_{\boldsymbol{r}}} \|v(\boldsymbol{x},t) - G(u_{\boldsymbol{r}}(\boldsymbol{x},t))\|^2_{L^2_\mu(V)}\,. \tag{48}$$

subject to the DO constraints

$$\left\langle \dot{\boldsymbol{\Psi}}_i^{\mathrm{T}} \boldsymbol{\Psi}_i \right\rangle_i = \mathbf{0}_{r_i\times r_i}, \qquad i = 1,\ldots,d-1, \tag{49}$$

which ensures that $\left\langle \boldsymbol{\Psi}_i^{\mathrm{T}}(t)\boldsymbol{\Psi}_i(t)\right\rangle_i = \boldsymbol{I}_{r_i\times r_i}$ for all $i = 1,\ldots,d-1$ and for all $t \geq 0$.

DO-TT propagator. It was shown in [8] that under these constraints, the convex minimization problem (48) admits a unique minimum for vectors in the tangent space (46) satisfying the PDE system

$$\begin{cases} \dot{\boldsymbol{\Psi}}_1 = \left[\left\langle G(u_{\boldsymbol{r}})\boldsymbol{\Psi}^{\mathrm{T}}_{\geq 2}\right\rangle_{\geq 2} - \boldsymbol{\Psi}_1 \left\langle \boldsymbol{\Psi}_1^{\mathrm{T}} G(u_{\boldsymbol{r}})\boldsymbol{\Psi}^{\mathrm{T}}_{\geq 2}\right\rangle_{\geq 1}\right] \left\langle \boldsymbol{\Psi}_{\geq 2}\boldsymbol{\Psi}^{\mathrm{T}}_{\geq 2}\right\rangle^{-1}_{\geq 2}, \\ \dot{\boldsymbol{\Psi}}_k = \Big[\left\langle \boldsymbol{\Psi}^{\mathrm{T}}_{\leq k-1} G(u_{\boldsymbol{r}})\boldsymbol{\Psi}^{\mathrm{T}}_{\geq k+1}\right\rangle_{\leq k-1,\geq k+1} - \\ \qquad\qquad \boldsymbol{\Psi}_k \left\langle \boldsymbol{\Psi}^{\mathrm{T}}_{\leq k} G(u_{\boldsymbol{r}})\boldsymbol{\Psi}^{\mathrm{T}}_{\geq k+1}\right\rangle_{\geq 1}\Big] \left\langle \boldsymbol{\Psi}_{\geq k+1}\boldsymbol{\Psi}^{\mathrm{T}}_{\geq k+1}\right\rangle^{-1}_{\geq k+1}, \qquad k = 2,3,\ldots,d-1, \\ \dot{\boldsymbol{\Psi}}_d = \left\langle \boldsymbol{\Psi}^{\mathrm{T}}_{\leq d-1} G(u_{\boldsymbol{r}})\right\rangle_{\leq d-1}. \end{cases} \tag{50}$$

Here, $u_{\boldsymbol{r}}(\boldsymbol{x},t)=\boldsymbol{\Psi}_1(t)\boldsymbol{\Psi}_2(t)\cdots\boldsymbol{\Psi}_d(t)\in\mathcal{M}_r$ and we have introduced the notation

$$\begin{aligned}\langle\boldsymbol{\Psi}\rangle_{\leq k}&=\int_{V_1\times\cdots\times V_k}\boldsymbol{\Psi}(\boldsymbol{x})d\mu_1(x_1)\cdots\mu_k(x_k),\\ \langle\boldsymbol{\Psi}\rangle_{\geq k}&=\int_{V_k\times\cdots\times V_d}\boldsymbol{\Psi}(\boldsymbol{x})d\mu_k(x_k)\cdots\mu_d(x_d),\\ \langle\boldsymbol{\Psi}\rangle_{\leq k-1,\geq k+1}&=\int_{V_1\times\cdots\times V_{k-1}\times V_{k+1}\times\cdots\times V_d}\boldsymbol{\Psi}(\boldsymbol{x})d\mu_1(x_1)\cdots\mu_{k-1}(x_{k-1})\mu_{k+1}(x_{k+1})\cdots\mu_d(x_d),\end{aligned}\tag{51}$$

for any matrix $\boldsymbol{\Psi}(\boldsymbol{x})\in M_{r\times s}\left(L^2_\mu(V)\right)$. The DO-FTT system (50) involves several inverse covariance matrices $\left\langle\boldsymbol{\Psi}_{\geq k}\boldsymbol{\Psi}^{\mathrm{T}}_{\geq k}\right\rangle^{-1}_{\geq k}$, which can become poorly conditioned in the presence of tensor modes with small energy (i.e. autocovariance matrices with small singular values). This phenomenon has been shown to be a result of the fact that the curvature of the tensor manifold at a tensor is inversely proportional to the smallest singular value present in the tensor [10, section 4]. To overcome the problem of inverting potentially ill-conditioned covariance matrices a rank-adaptive operator splitting method was proposed in [6].

Numerical application of DO-TT to the Fokker-Planck equation. We have seen in Chapter 2 that the Fokker–Planck equation describes the evolution of the probability density function (PDF) of the state vector solving the Itô stochastic differential equation (SDE) [13]

$$d\boldsymbol{X}_t=\boldsymbol{\mu}(\boldsymbol{X}_t,t)dt+\boldsymbol{\sigma}(\boldsymbol{X}_t,t)d\boldsymbol{W}_t.\tag{52}$$

Here, $\boldsymbol{X}_t$ is the d-dimensional state vector, $\boldsymbol{\mu}(\boldsymbol{X}_t,t)$ is the d-dimensional drift, $\boldsymbol{\sigma}(\boldsymbol{X}_t,t)$ is an $d\times m$ matrix and $\boldsymbol{W}_t$ is an m-dimensional standard Wiener process. The Fokker–Planck equation that corresponds to (52) has the form

$$\begin{cases}\dfrac{\partial p(\boldsymbol{x},t)}{\partial t}=\mathcal{L}(\boldsymbol{x},t)p(\boldsymbol{x},t),\\[2ex] p(\boldsymbol{x},0)=p_0(\boldsymbol{x}),\end{cases}\tag{53}$$

where $p_0(\boldsymbol{x})$ is the PDF of the initial state $\boldsymbol{X}_0$, $\mathcal{L}$ is a second-order linear differential operator defined as

$$\mathcal{L}(\boldsymbol{x},t)p(\boldsymbol{x},t)=-\sum_{k=1}^{d}\frac{\partial}{\partial x_k}\left(\mu_k(x,t)p(\boldsymbol{x},t)\right)+\sum_{k,j=1}^{d}\frac{\partial^2}{\partial x_k\partial x_j}\left(D_{ij}(\boldsymbol{x},t)p(\boldsymbol{x},t)\right),\tag{54}$$

and $\boldsymbol{D}(\boldsymbol{x},t)=\boldsymbol{\sigma}(\boldsymbol{x},t)\boldsymbol{\sigma}(\boldsymbol{x},t)^{\mathrm{T}}/2$ is the diffusion tensor. For our numerical demonstration we set

$$\boldsymbol{\mu}(\boldsymbol{x})=\alpha\begin{bmatrix}\sin(x_1)\\ \sin(x_3)\\ \sin(x_4)\\ \sin(x_1)\end{bmatrix},\qquad\boldsymbol{\sigma}(\boldsymbol{x})=\sqrt{2\beta}\begin{bmatrix}g(x_2)&0&0&0\\ 0&g(x_3)&0&0\\ 0&0&g(x_4)&0\\ 0&0&0&g(x_1)\end{bmatrix},\tag{55}$$

where $g(x)=\sqrt{1+k\sin(x)}$. With the drift and diffusion matrices chosen in (55) the operator (54) takes the form

$$\begin{aligned}\mathcal{L}=&-\alpha\left(\cos(x_1)+\sin(x_1)\frac{\partial}{\partial x_1}+\sin(x_3)\frac{\partial}{\partial x_2}+\sin(x_4)\frac{\partial}{\partial x_3}+\sin(x_1)\frac{\partial}{\partial x_4}\right)\\ &+\beta\left((1+k\sin(x_2))\frac{\partial^2}{\partial x_1^2}+(1+k\sin(x_3))\frac{\partial^2}{\partial x_2^2}+(1+k\sin(x_4))\frac{\partial^2}{\partial x_3^2}+(1+k\sin(x_1))\frac{\partial^2}{\partial x_4^2}\right).\end{aligned}\tag{56}$$

Clearly $\mathcal{L}$ is a linear, time-independent separable operator of rank 9, since it can be written as

$$\mathcal{L} = \sum_{i=1}^{9} L_i^{(1)} \otimes L_i^{(2)} \otimes L_i^{(3)} \otimes L_i^{(4)}, \tag{57}$$

where each $L_i^{(j)}$ operates on x_j only. Specifically, we have

$$\begin{aligned}
&L_1^{(1)} = -\alpha\cos(x_1), \quad L_2^{(1)} = -\alpha\sin(x_1)\frac{\partial}{\partial x_1}, \quad L_3^{(2)} = -\alpha\frac{\partial}{\partial x_2}, \quad L_3^{(3)} = \sin(x_3),\\
&L_4^{(3)} = -\alpha\frac{\partial}{\partial x_3}, \quad L_4^{(4)} = \sin(x_4), \quad L_5^{(1)} = -\alpha\sin(x_1), \quad L_5^{(4)} = \frac{\partial}{\partial x_4},\\
&L_6^{(1)} = \beta\frac{\partial^2}{\partial x_1^2}, \quad L_6^{(2)} = 1 + k\sin(x_2), \quad L_7^{(2)} = \beta\frac{\partial^2}{\partial x_2^2}, \quad L_7^{(3)} = 1 + k\sin(x_3),\\
&L_8^{(3)} = \beta\frac{\partial^2}{\partial x_3^2}, \quad L_8^{(2)} = 1 + k\sin(x_4), \quad L_9^{(4)} = \beta\frac{\partial^2}{\partial x_4^2}, \quad L_9^{(1)} = 1 + k\sin(x_1),
\end{aligned} \tag{58}$$

and all other unspecified $L_i^{(j)}$ are identity operators. We set the parameters in (55) as $\alpha = 0.1$, $\beta = 2.0$, $k = 1.0$ and solve (53) on the four-dimensional flat torus $\mathbb{T}^4$. The initial PDF is set as

$$p_0(\boldsymbol{x}) = \frac{\sin(x_1)\sin(x_2)\sin(x_3)\sin(x_4) + 1}{16\pi^4}. \tag{59}$$

Note that (59) is a four-dimensional FTT tensor with multilinear rank $\boldsymbol{r} = \begin{bmatrix} 1 & 2 & 2 & 2 & 1 \end{bmatrix}$. Upon normalizing the modes appropriately we obtain the left orthogonalized initial condition required to begin integration

$$\begin{aligned}
p_0(\boldsymbol{x}) = &\psi_1(1; x_1; 1)\psi_2(1; x_2; 1)\psi_3(1; x_3; 1)\psi_4(1; x_4; 1)\sqrt{\lambda(1)}\\
&+ \psi_1(1; x_1; 2)\psi_2(2; x_2; 2)\psi_3(2; x_3; 2)\psi_4(2; x_4; 1)\sqrt{\lambda(2)},
\end{aligned} \tag{60}$$

where

$$\psi_i(1; x_i; 1) = \frac{\sin(x_i)}{\sqrt{\pi}}, \qquad \sqrt{\lambda(1)} = \frac{1}{16\pi^2}. \tag{61}$$

All other tensor modes are equal to $1/\sqrt{2\pi}$, and $\sqrt{\lambda(2)} = 1/(2\pi^2)$. To obtain a benchmark solution with which to compare the rank-adaptive FTT solution, we solve the PDE (53) using a Fourier pseudo-spectral method on the flat torus $\mathbb{T}^4$ with $21^4 = 194481$ evenly-spaced points. As before, the operator $\mathcal{L}$ is represented in terms of pseudo-spectral differentiation matrices, and the resulting semi-discrete approximation (ODE system) is integrated with an explicit fourth-order Runge Kutta method using time step $\Delta t = 10^{-4}$. The numerical solution we obtained in this way is denoted by $p_{\text{ref}}(\boldsymbol{x}, t)$. We also solve the Fokker-Planck using the proposed rank-adaptive FTT method with first-order Lie-Trotter time integrator and normal vector thresholding. We run three simulations all with time step $\Delta t = 10^{-4}$: one with no rank adaption, and two with rank-adaptation and normal component thresholds set to $\epsilon_{\text{inc}} = 10^{-3}$ and $\epsilon_{\text{inc}} = 10^{-4}$. In Figure 5 we plot three time snapshots of the two-dimensional solution marginal

$$p(x_1, x_2, t) = \int_0^{2\pi}\int_0^{2\pi} p(x_1, x_2, x_3, x_4, t)dx_3dx_4 \tag{62}$$

computed with the rank-adaptive FTT integrator ($\epsilon_{\text{inc}} = 10^{-4}$) and the full tensor product pseudo-spectral method (reference solution). In Figure 6(a) we compare the $L^2(\Omega)$ errors of the rank-adaptive method relative to the reference solution. It is seen that as we decrease the threshold the solution becomes more accurate. In Figure 6(b) we plot the component of $\mathcal{L}p_{\boldsymbol{r}}$ normal to the tensor manifold. Note that in the rank-adaptive FTT solution with thresholds $\epsilon_{\text{inc}} = 10^{-3}$ and $\epsilon_{\text{inc}} = 10^{-4}$ the solver performs both mode addition as well as mode removal. This is documented in Figure 7. The abrupt change in rank

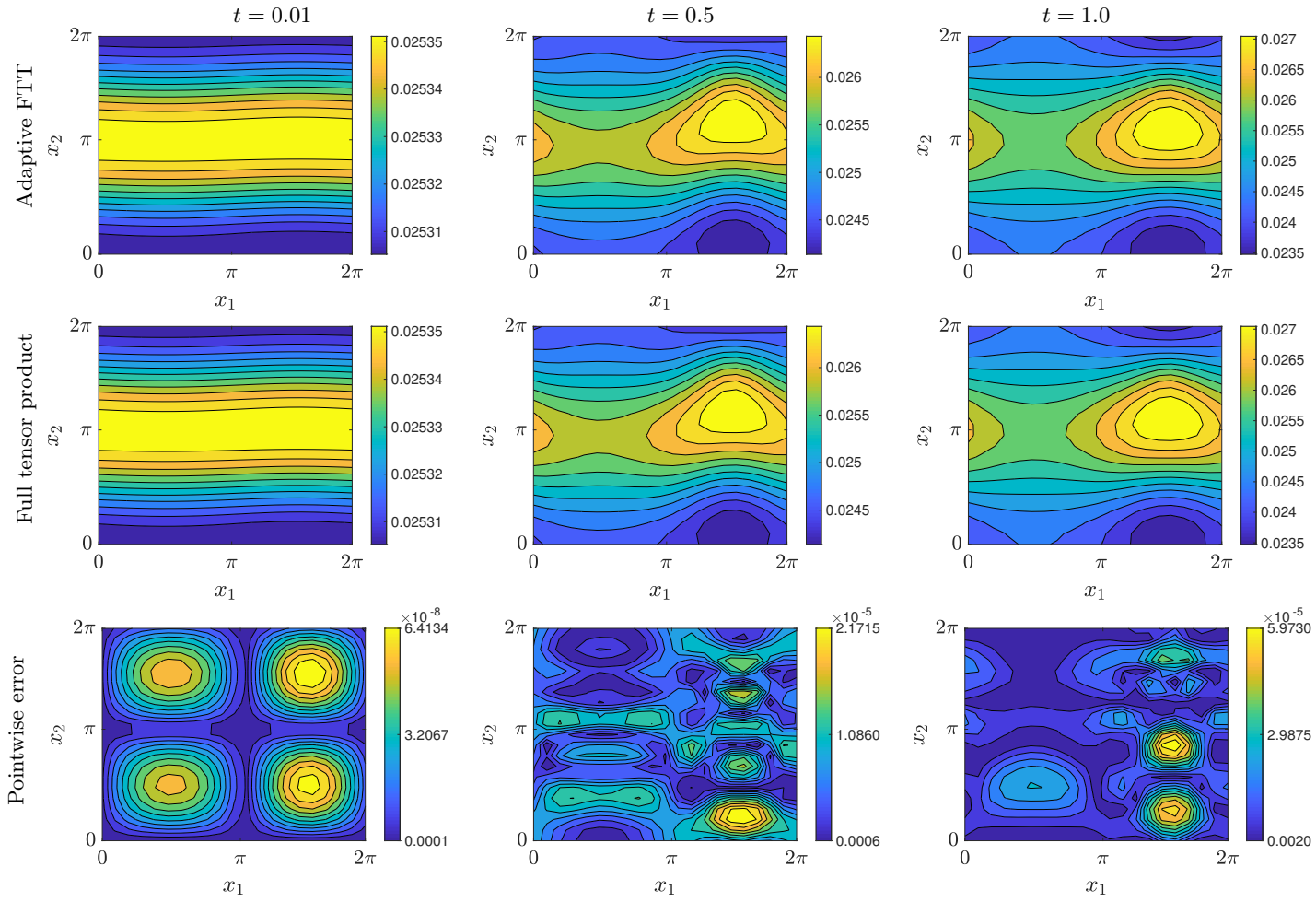

Figure 5: Time snapshots of marginal PDF $p_{\boldsymbol{r}}(x_1, x_2, t)$ corresponding to the solution to the Fokker-Planck equation (53). We plot marginals computed with the rank-adaptive FTT integrator using $\epsilon_{\text{inc}} = 10^{-4}$ (top row) and with the full tensor product Fourier pseudo-spectral method (middle row). We also plot the pointwise error between the two numerical solutions (bottom row). The initial condition is the FTT tensor (59).

observed in Figure 7(a)-(c) near time $t = 0.4$ corresponding to the rank-adaptive solution with threshold $\epsilon\text{inc} = 10^{-4}$ is due to the time step size Δt being equal to ϵ_{inc}. This can be justified as follows. Recall that the solution is first order accurate in Δt and therefore the approximation of the component of $\mathcal{L}p_{\boldsymbol{r}}$ normal to the tensor manifold $\mathcal{M}_{\boldsymbol{r}}$ is first-order accurate in Δt. If we set $\epsilon_{\text{inc}} \leq \Delta t$, then the rank-adaptive scheme may overestimate the number of modes needed to achieve accuracy on the order of Δt. This does not affect the accuracy of the numerical solution due to the robustness of the Lie-Trotter integrator to over-approximation [11]. Moreover we notice that the rank-adaptive scheme removes the unnecessary modes ensure that the tensor rank is not unnecessarily large. In fact, the diffusive nature of the Fokker-Plank equation on the flat torus $\mathbb{T}^4$ yields relaxation to a statistical equilibrium state that depends on the drift and diffusion coefficients in (53). Such an equilibrium state may be well-approximated by a low-rank FTT tensor.

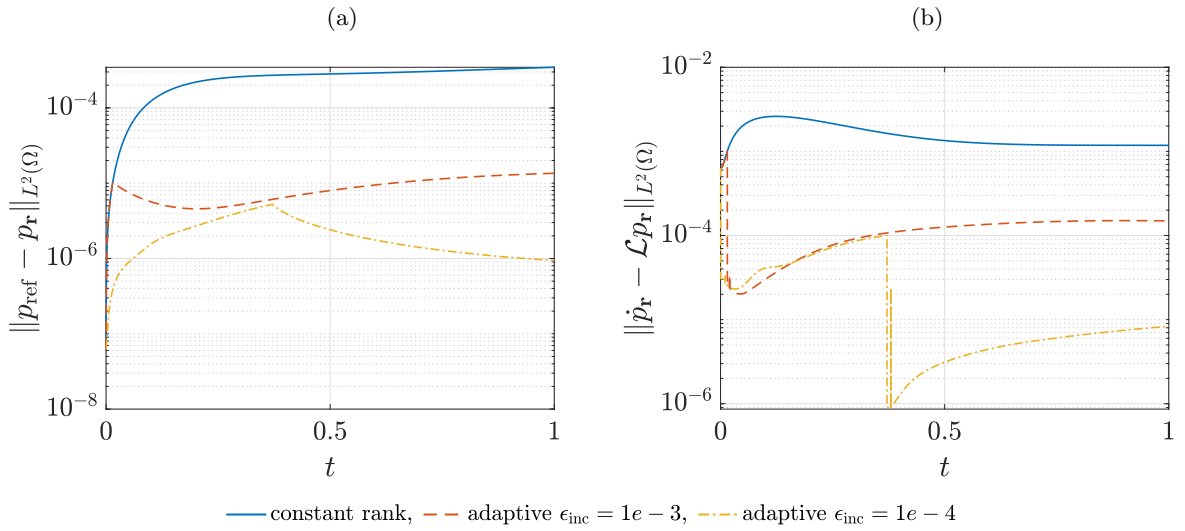

Figure 6: (a) The $L^2(\Omega)$ error of the FTT solution $p_{\boldsymbol{r}}(\boldsymbol{x}, t)$ relative to the benchmark solution $p_{\text{ref}}(\boldsymbol{x}, t)$ computed with a Fourier pseudo-spectral method on a tensor product grid. (b) Norm of the component of $\mathcal{L}p_{\boldsymbol{r}}$ normal to the tensor manifold (see Figure 4). Such component is approximated a two-point BDF formula at each time step.

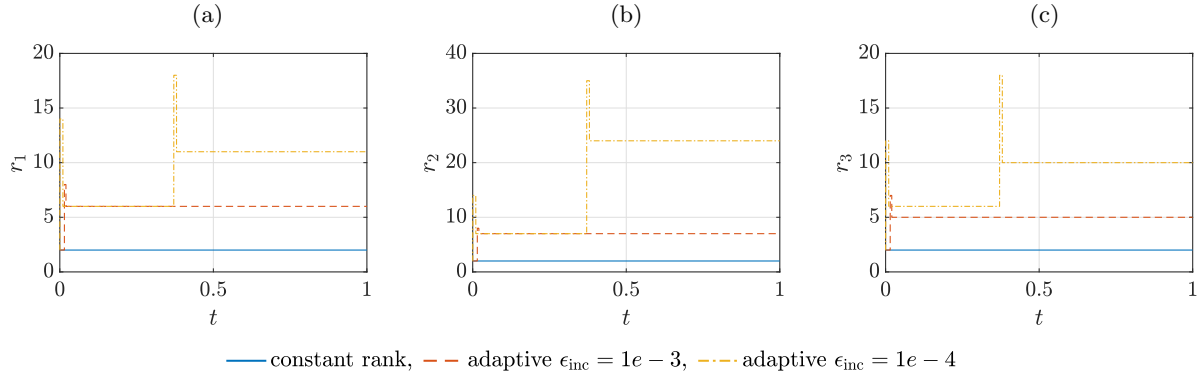

Figure 7: Tensor rank $\boldsymbol{r} = [1\, r_1\, r_2\, r_3\, 1]$ of adaptive FTT solution to the four dimensional Fokker-Planck equation (53).

References

[1] H. Babaee, M. Choi, T. P. Sapsis, and G. E. Karniadakis. A robust bi-orthogonal/dynamically-orthogonal method using the covariance pseudo-inverse with application to stochastic flow problems. *J. Comput. Phys.*, 344:303–319, 2017.

[2] D. Bigoni, A. P. Engsig-Karup, and Y. M. Marzouk. Spectral tensor-train decomposition. *SIAM J. Sci. Comput.*, 38(4):A2405–A2439, 2016.

[3] M. Cheng, T. Y. Hou, and Z. Zhang. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations I: derivation and algorithms. *J. Comput. Phys.*, 242:843–868, 2013.

[4] M. Cheng, T. Y. Hou, and Z. Zhang. A dynamically bi-orthogonal method for time-dependent stochastic partial differential equations II: adaptivity and generalizations. *J. Comput. Phys.*, 242:753–776, 2013.

[5] M. Choi, T. Sapsis, and G. E. Karniadakis. On the equivalence of dynamically orthogonal and bi-orthogonal methods: Theory and numerical simulations. *J. Comput. Phys.*, 270:1–20, 2014.

[6] A. Dektor, A. Rodgers, and D. Venturi. Rank-adaptive tensor methods for high-dimensional nonlinear pdes. *Journal of Scientific Computing*, 88(36):1–27, 2021.

[7] A. Dektor and D. Venturi. Dynamically orthogonal tensor methods for high-dimensional nonlinear PDEs. *J. Comput. Phys.*, 404:109125, 2020.

[8] A. Dektor and D. Venturi. Dynamic tensor approximation of high-dimensional nonlinear pdes. *J. Comput. Phys.*, 437:110295, 2021.

[9] M. Gerritsma, J.-B. van der Steen, P. Vos, and G. E. Karniadakis. Time-dependent generalized polynomial chaos. *J. Comput. Phys.*, 229(22):8333–8363, 2010.

[10] O. Koch and C. Lubich. Dynamical low-rank approximation. *SIAM J. Matrix Anal. Appl.*, 29(2):434–454, 2007.

[11] C. Lubich, I. V. Oseledets, and B. Vandereycken. Time integration of tensor trains. *SIAM J. Numer. Anal.*, 53(2):917–941, 2015.

[12] I. V. Oseledets. Constructive representation of functions in low-rank tensor formats. *Constructive Approximation*, 37:1–18, 2013.

[13] H. Risken. *The Fokker-Planck equation: methods of solution and applications*. Springer-Verlag, second edition, 1989. Mathematics in science and engineering, vol. 60.

[14] T. P. Sapsis and P. F .J. Lermusiaux. Dynamically orthogonal field equations for continuous stochastic dynamical systems. *Physica D*, 238(23-24):2347–2360, 2009.

[15] A. Uschmajew and B. Vandereycken. The geometry of algorithms using hierarchical tensors. *Linear Algebra Appl.*, 439(1):133–166, 2013.