Time-dependent Hermite–Galerkin spectral method and its applications

Xue Luo[a], Shing-Tung Yau[b], Stephen S.-T. Yau[c,*]

[a] School of Mathematics and System Sciences, Beihang University, Beijing 100191, PR China
[b] Department of Mathematics, Harvard University, Cambridge, MA 02138, USA
[c] Department of Mathematical Sciences, Tsinghua University, Beijing 100084, PR China

ARTICLE INFO

ABSTRACT

A time-dependent Hermite–Galerkin spectral method (THGSM) is investigated in this paper for the nonlinear convection–diffusion equations in the unbounded domains. The time-dependent scaling factor and translating factor are introduced in the definition of the generalized Hermite functions (GHF). As a consequence, the THGSM based on these GHF has many advantages, not only in theoretical proofs, but also in numerical implementations. The stability and spectral convergence of our proposed method have been established in this paper. The Korteweg–de Vries–Burgers (KdVB) equation and its special cases, including the heat equation and the Burgers' equation, as the examples, have been numerically solved by our method. The numerical results are presented, and it surpasses the existing methods in accuracy. Our theoretical proof of the spectral convergence has been supported by the numerical results.

1. Introduction

Many scientific and engineering problems are naturally modeled in the unbounded domains. One way to numerically solve the problems is to restrict the model equation in some bounded domain and artificially impose some boundary condition cleverly. Whereas this introduces errors even before the implementation of the numerical scheme. Another more suitable way is to use the spectral approaches employing orthogonal systems in unbounded domain, such as using Laguerre polynomials for the problems in semi-bounded or exterior domains [13,19], and using Hermite polynomials for those in the whole space [1,10,11,15,22].

Although the freedom from artificial boundary condition is very attractive, the Hermite spectral method (HSM) is only widely studied in the recent decade, due to its poor resolution without the appropriate scaling factor. The Hermite functions, defined as $\{H_n(x)e^{-x^2}\}_{n=0}^{\infty}$, have the same deficiency as the polynomials $\{H_n(x)\}_{n=0}^{\infty}$. However, a suitably chosen scaling factor will greatly improve the resolution. Its importance has been discussed in [25,28]. It has been shown in [4] that the scaling factor should be selected according to the truncated modes N and the asymptotical behavior of the function $f(x)$, as $|x| \to \pm\infty$. The optimal scaling factor is still an open problem, even in the case that $f(x)$ is given explicitly, to say nothing of the exact solution to a differential equation, which is in general unknown beforehand. Recently, during the study of using the HSM to solve the nonlinear filtering problems, the first and the third author of this paper gave a practical strategy in [21] to pick the appropriate scaling factor and the corresponding truncated mode for at least the most commonly used types of functions, i.e. the Gaussian type and the super-Gaussian type. Thanks to this guideline, the Hermite–Galerkin spectral method (HGSM) becomes implementable.

---

* Corresponding author. Tel.: +86 10 62787874.
  E-mail addresses: xluo@buaa.edu.cn (X. Luo), yau@math.harvard.edu (S.-T. Yau), yau@uic.edu (S.S.-T. Yau).

In the literature of solving partial differential equations in unbounded domains using HSM, nearly all the schemes are not direct Galerkin ones. As far as the author knows, there are at least two possible reasons:

1. The lack of the practical guidelines of choosing appropriate scaling factor makes the direct Galerkin method infeasible.
2. When directly applying traditional definition of Hermite functions (i.e., $\{H_n(x)e^{-x^2}\}_{n=0}^\infty$) to second-order differential equations, it is found in [12] that the stiff matrix is of nonsymmetric bilinear form, which has no property of coerciveness. In other word, the stability cannot be established by using the classical energy method.

To overcome the obstacles above, Funaro and Kavian [11] first consider the use of the Hermite polynomials to approximate the solutions of some diffusion evolution equations in unbounded domains. The variable transformation technique is introduced to get better resolution. Later, Guo [12], Guo and Xu [15] developed the Hermite polynomial spectral and pseudo-spectral methods, where the transformation $U = e^{-x^2}V$ is used, and then V is approximated by the Hermite polynomials. Ma et al. [22], Ma and Zhao [23] introduced a time-dependent parameter to stabilize the scheme, which is based on the traditional defined Hermite functions. However, no discussion was given on how to choose such parameter for the particular problems.

The aim of this paper is to develop a time-dependent Hermite–Galerkin spectral method (THGSM) to approximate the solution to the nonlinear convection–diffusion equations with high accuracy. In this paper, we focus on the nonlinearity satisfying *Assumptions 1 and 2* in Section 3, which includes the Burger's equation and KdVB equation. The spectral method has been applied to various nonlinear problems, such as the nonlinear Shröndinger equation [2], Bose–Einstein condensates [3], Navier–Stokes equation [20] etc. Also different types of differential equations have been studied by using spectral method, for instance, pantography-type differential and integral equations [5], hyperbolic PDEs [6], pattern-forming nonlinear evolution systems [7] etc. More thorough discussion of spectral methods is referred to [26]. The time-dependence is reflected in the definition of the generalized Hermite functions (GHF), where the scaling factor and the translating factor are the functions of time. The choice of the time-dependent scaling factor can follow the guidelines in [21], while the time-dependent translating factor mainly deals with the time-shifting of the solution, see examples in Section 4.3. The advantages of our THGSM are the following:

1. It is a direct Galerkin scheme, which can be implemented straightforward. And the resulting stiffness matrix of the second-order differential equations is of nice properties. For example, it is tri-diagonal, symmetric and diagonally dominant in the linear case, i.e. $g(u) \equiv 0$ in (3.1); it is symmetric for the Burgers' equation, i.e. $g(u) = \frac{u}{2}$ in (3.1).
2. The proofs of stability and spectral convergence are greatly simplified, thanks to the definition of GHF. They are analyzed in the L^2 space, instead of the weighted one as in [22].
3. From the numerical simulations in Section 4, our scheme outperforms nearly all the existing methods in accuracy.

An outline of the paper is as follows. In Section 2, we give the definition of GHF and its properties. For the readers' convenience, we include the proof of the error estimate of the orthogonal projection. Our TGHSM to solve the nonlinear convection–diffusion equations is introduced in Section 3. The stability analysis in the sense of [12] and the spectral convergence are shown there. Section 4 is devoted to the numerical simulations, where we compared the numerical results with those obtained by other methods in some benchmark equations.

2. Generalized Hermite functions (GHF)

We introduce the GHF and derive some properties which are inherited from the physical Hermite polynomials. For the sake of completeness, we give the proof of the convergence rate of the orthogonal approximation.

2.1. Notations and preliminaries

Let $L^2(\mathbb{R})$ be the Lebesgue space, which equips with the norm $\|\cdot\| = (\int_\mathbb{R} |\cdot|^2 dx)^{\frac{1}{2}}$ and the scalar product $\langle \cdot, \cdot \rangle$.

Let $\mathcal{H}_n(x)$ be the physical Hermite polynomials given by $\mathcal{H}_n(x) = (-1)^n e^{x^2} \partial_x^n e^{-x^2}$, $n \geq 0$. The three-term recurrence

$$\mathcal{H}_0 \equiv 1; \quad \mathcal{H}_1(x) = 2x; \quad \text{and} \quad \mathcal{H}_{n+1}(x) = 2x\mathcal{H}_n(x) - 2n\mathcal{H}_{n-1}(x). \tag{2.1}$$

is more handy in implementation. One of the well-known and useful fact of Hermite polynomials is that they are mutually orthogonal with the weight $w(x) = e^{-x^2}$. We define the time-dependent GHF as

$$H_n^{\alpha,\beta}(x, t) = \left(\frac{\alpha(t)}{2^n n! \sqrt{\pi}} \right)^{\frac{1}{2}} \mathcal{H}_n[\alpha(t)(x - \beta(t))] e^{-\frac{1}{2}\alpha^2(t)[x - \beta(t)]^2}, \tag{2.2}$$

for $n \geq 0$, where $\alpha(t) > 0$, $\beta(t)$, for $t \in [0, T]$, are functions of time. For the conciseness of notation, let us denote $d(n) = \sqrt{\frac{n}{2}}$. And if no confusion will arise, in the sequel we omit the t in $\alpha(t)$ and $\beta(t)$. It is readily to derive the following properties for the GHF (2.2):

■ At each time $t > 0$, $\{H_n^{\alpha,\beta}(\cdot, t)\}_{n \in \mathbb{Z}^+}$ form the orthogonal basis of $L^2(\mathbb{R})$, i.e.

$$\int_\mathbb{R} H_n^{\alpha,\beta}(x, t) H_m^{\alpha,\beta}(x, t) dx = \delta_{nm}, \tag{2.3}$$

where δ_{nm} is the Kronecker function.

■ $H_n^{\alpha,\beta}(\cdot, t)$ is the nth eigenfunction of the following Strum–Liouville problem

$$\mathcal{L}^{\alpha,\beta} u(x, t) = \lambda_n(t)u(x, t),$$

where

$$\mathcal{L}^{\alpha,\beta}(\circ) = -e^{\frac{1}{2}\alpha^2(x-\beta)^2} \partial_x (e^{-\alpha^2(x-\beta)^2} \partial_x (e^{\frac{1}{2}\alpha^2(x-\beta)^2} \circ))$$

with the corresponding eigenvalue $\lambda_n(t) = 2\alpha^2(t)n$.

■ By convention, $H_n^{\alpha,\beta} \equiv 0$, for $n < 0$. For $n \geq 0$, the three-term recurrence is inherited from the Hermite polynomials:

$$xH_n^{\alpha,\beta}(x, t) = \frac{d(n+1)}{\alpha} H_{n+1}^{\alpha,\beta}(x, t) + \beta H_n^{\alpha,\beta}(x, t) + \frac{d(n)}{\alpha} H_{n-1}^{\alpha,\beta}(x, t);$$

or $\quad \alpha(x - \beta)H_n^{\alpha,\beta}(x, t) = d(n+1)H_{n+1}^{\alpha,\beta}(x, t) + d(n)H_{n-1}^{\alpha,\beta}(x, t).$ \hfill (2.4)

■ The derivative of $H_n^{\alpha,\beta}(x, t)$ with respect to x and t

$$\partial_x H_n^{\alpha,\beta}(x, t) = -\frac{1}{2}\sqrt{\lambda_{n+1}} H_{n+1}^{\alpha,\beta}(x, t) + \frac{1}{2}\sqrt{\lambda_n} H_{n-1}^{\alpha,\beta}(x, t)$$
$$= -d(n+1)\alpha H_{n+1}^{\alpha,\beta}(x, t) + d(n)\alpha H_{n-1}^{\alpha,\beta}(x, t). \hfill (2.5)$$

■ The "orthogonality" of $\left\{\partial_x H_n^{\alpha,\beta}(\cdot, t)\right\}_{n \in \mathbb{Z}^+}$

$$\int_{\mathbb{R}} \partial_x H_n^{\alpha,\beta} \partial_x H_m^{\alpha,\beta} dx = \begin{cases} \alpha^2 \left[d^2(n+1) + d^2(n)\right], & \text{if } m = n; \\ -\alpha^2 d(l+1)d(l+2), & l = \min\{n, m\}, \text{ if } |n - m| = 2; \\ 0, & \text{otherwise.} \end{cases} \hfill (2.6)$$

Lemma 2.1 (Derivative with respect to. $t)$

$$\partial_t H_n^{\alpha,\beta}(x, t) = -\frac{\alpha'}{\alpha} d(n+1)d(n+2)H_{n+2}^{\alpha,\beta}(x, t) + \alpha\beta' d(n+1)H_{n+1}^{\alpha,\beta}$$
$$- \alpha\beta' d(n)H_{n-1}^{\alpha,\beta}(x, t) + \frac{\alpha'}{\alpha} d(n)d(n-1)H_{n-2}^{\alpha,\beta}(x, t), \hfill (2.7)$$

where $d(n) = \sqrt{\frac{n}{2}}$ and α, β are the functions of t.

Proof. Through direct computations, we have

$$\partial_t H_n^{\alpha,\beta}(x, t) \overset{(2.2)}{=} \frac{\alpha'}{2\alpha} H_n^{\alpha,\beta}(x, t) + \sqrt{2n} H_{n-1}^{\alpha,\beta}(x, t)[\alpha'(x - \beta) - \alpha\beta'] - [\alpha'(x - \beta) - \alpha\beta']\alpha(x - \beta)H_n^{\alpha,\beta}(x, t)$$

$$\overset{(2.4)}{=} \frac{\alpha'}{2\alpha} H_n^{\alpha,\beta}(x, t) + \sqrt{n(n-1)}\frac{\alpha'}{\alpha} H_{n-2}^{\alpha,\beta}(x, t) - \alpha\beta'\sqrt{2n} H_{n-1}^{\alpha,\beta}(x, t) + n\frac{\alpha'}{\alpha} H_n^{\alpha,\beta}(x, t)$$

$$\quad - \frac{\alpha'}{2\alpha}\sqrt{n(n-1)} H_{n-2}^{\alpha,\beta}(x, t) + \alpha\beta'\sqrt{\frac{n}{2}} H_{n-1}^{\alpha,\beta}(x, t) - \frac{\alpha'}{2\alpha}(2n+1)H_n^{\alpha,\beta}(x, t)$$

$$\quad + \alpha\beta'\sqrt{\frac{n+1}{2}} H_{n+1}^{\alpha,\beta}(x, t) - \frac{\alpha'}{2\alpha}\sqrt{(n+1)(n+2)} H_{n+2}^{\alpha,\beta}(x, t)$$

$$= \frac{\alpha'}{2\alpha}\sqrt{n(n-1)} H_{n-2}^{\alpha,\beta}(x, t) - \alpha\beta'\sqrt{\frac{n}{2}} H_{n-1}^{\alpha,\beta}(x, t) + \alpha\beta'\sqrt{\frac{n+1}{2}} H_{n+1}^{\alpha,\beta}(x, t) - \frac{\alpha'}{2\alpha}\sqrt{(n+1)(n+2)} H_{n+2}^{\alpha,\beta}(x, t).$$

$\square$

Remark 2.1. The simple fact follows immediately from (2.7) and (2.1). For $N \gg 1$ sufficiently large, for $t \in [0, T]$, then

$$\left\|\left|\partial_t H_N^{\alpha,\beta}(x, t)\right|\right\|^2 \lesssim \frac{\alpha'^2}{\alpha^2} N^2 + \alpha'\beta' N^{\frac{3}{2}} + \alpha^2\beta'^2 N.$$

In particular, we have

$$\left\|\left|\partial_t H_N^{\alpha,\beta}(x, t)\right|\right\|^2 \lesssim \frac{\alpha'^2}{\alpha^2} N^2.$$

We follow the convection in the asymptotic analysis, $a \sim b$ means that there exists some constants $C_1, C_2 > 0$ such that $C_1 a \leq b \leq C_2 a$; $a \lesssim b$ means that there exists some constant $C_3 > 0$ such that $a \leq C_3 b$.

Any function $u(\cdot, t) \in L^2(\mathbb{R})$ can be written in the form

$$u(\cdot, t) = \sum_{n=0}^{\infty} \hat{u}_n(t) H_n^{\alpha,\beta}(\cdot, t), \quad \hat{u}_n(t) = \int_{\mathbb{R}} u(x, t) H_n^{\alpha,\beta}(x, t)dx := \left\langle u, H_n^{\alpha,\beta}\right\rangle. \hfill (2.8)$$

where $\{\hat{u}_n\}_{n=0}^{\infty}$ are the Fourier–Hermite coefficients defined similarly as Fourier coefficients with $H_n^{\alpha,\beta}(x,t)$ taking the place of harmonic oscillators.

For $N \geq 0$, let

$$\mathcal{R}_N(t) = \text{span}\left\{H_0^{\alpha,\beta}(x,t), \ldots, H_N^{\alpha,\beta}(x,t)\right\}.$$

At each time $t \in [0, T]$, it is a linear subspace of $L^2(\mathbb{R})$.

Lemma 2.2. *For any function* $\varphi(x,t) \in \mathcal{R}_N(t)$, *we have*

$$||\partial_x\varphi|| \lesssim \alpha N^{\frac{1}{2}}||\varphi||.$$

Proof. For any function $\varphi(x,t) \in \mathcal{R}_N(t)$, we can write it as

$$\varphi(x,t) = \sum_{k=0}^{N}\hat{\varphi}_k(t)H_k^{\alpha,\beta}(x,t).$$

Thus,

$$
\begin{aligned}
||\partial_x\varphi||^2 &= \langle \partial_x\varphi, \partial_x\varphi \rangle \\
&= \alpha^2 \sum_{k=0}^{N}\hat{\varphi}_k\left\langle -d(k+1)H_{k+1}^{\alpha,\beta} + d(k)H_{k-1}^{\alpha,\beta}, \sum_{l=0}^{N}\hat{\varphi}_l\left[-d(l+1)H_{l+1}^{\alpha,\beta} + d(l)H_{l-1}^{\alpha,\beta}\right]\right\rangle \\
&= \sum_{k=0}^{N}\hat{\varphi}_k^2[d^2(k+1) + d^2(k)] - 2\alpha^2\sum_{k=2}^{N}d(k-1)d(k)\hat{\varphi}_{k-2}\hat{\varphi}_k \lesssim \alpha^2 N\sum_{k=0}^{N}\hat{\varphi}_k^2 \leq \alpha^2 N||\varphi||^2. \quad \square
\end{aligned}
$$

2.2. Orthogonal projection and approximations

Let us define the norm of the space $H_{r,\alpha}(\mathbb{R})$ as below:

Definition 2.1. For any integer $r \geq 0$,

$$H_{r,\alpha}(\mathbb{R}) := \left\{u \in L^2(\mathbb{R}) : ||u||_r < \infty, ||u||_r^2 := \sum_{k=0}^{\infty}\lambda_{k+1}^r\hat{u}_k^2, \lambda_k = 2\alpha^2 k\right\}. \tag{2.9}$$

It is readily shown in [31,32] that when $\beta = 0$ the norm defined in (2.9) is equivalent to the following Sobolev-like norm

$$||u||_{P,r,\alpha}^2 = \sum_{k=0}^{r}\left|\left|(\alpha^4 x^2 + \alpha^2)^{\frac{r-k}{2}}\partial_x^k u\right|\right|^2$$

in the sense that for any $u \in H_{r,\alpha}(\mathbb{R})$, with $r \geq 0$

$$||u||_{r,\alpha}^2 \sim \alpha||u||_{P,r,\alpha}^2.$$

This equivalence can be extended trivially to where $\beta \neq 0$, by mimicking the proof of Lemmas 2.2–2.3 in [31] with x replaced by $(x - \beta)$.

From the definition of $||\cdot||_{P,r,\alpha}$, it is clear to see that the larger the r is, the smaller space $H_{r,\alpha}$ is, and the smoother the functions are. The index r can be viewed as the indicator of the regularity of the functions.

Next we need some estimates on $||x^\gamma \partial_x^s u(x)||^2$, for any integer $\gamma, s \geq 0$:

Lemma 2.3. *For any function* $u \in H_{s,\alpha}(\mathbb{R})$, *with some integer* $s \geq 0$, *we have*

$$||\partial_x^s u|| \lesssim ||u||_{s,\alpha}. \tag{2.10}$$

Proof. For any integer $s \geq 0$,

$$||\partial_x^s u||^2 = \left|\left|\sum_{n=0}^{\infty}\hat{u}_n\partial_x^s H_n^{\alpha,\beta}\right|\right|^2 = \left|\left|\sum_{n=0}^{\infty}\hat{u}_n\sum_{k=-s}^{s}a_{n,k}H_{n+k}^{\alpha,\beta}\right|\right|^2,$$

where for each n fixed, $a_{n,k}$ is a product of s factors of $\sqrt{\lambda_{n+j}}$ with $-s \leq j \leq s$. Notice that $\lambda_{n+j} \sim \lambda_{n+1}$ for $n + j \geq 0$ and $H_{n+j}^{\alpha,\beta} \equiv 0$ for $n + j < 0$. Hence, we have

$$||\partial_x^s u||^2 \overset{(2.3)}{\lesssim} \sum_{n=0}^{\infty} \lambda_{n+1}^s \hat{u}_n^2 = ||u||_{s,\alpha}^2,$$

for any integer $s \geq 0$. $\square$

In the spectral method, a function in $L^2(\mathbb{R})$ is approximated by the partial sum of the first N frequency modes. We define the L^2-orthogonal projection $P_N^{\alpha,\beta} : L^2(\mathbb{R}) \to \mathcal{R}_N(t)$, for some $t \in [0, T]$. Given $v \in L^2(\mathbb{R})$,

$$\left\langle v - P_N^{\alpha,\beta} v, \phi \right\rangle = 0, \quad \forall \phi \in \mathcal{R}_N(t).$$

More precisely,

$$P_N v(x) := \sum_{n=0}^{N} \hat{v}_n(t) H_n^{\alpha,\beta}(x, t),$$

for some $t \in [0, T]$, where $\hat{v}_n$ are the Fourier–Hermite coefficients defined in (2.8).

And the truncated error $||u - P_N u||_{P, r, \alpha}$, for any integer $r \geq 0$, has been estimated readily in Theorem 2.3, [14] for $\alpha = 1$, $\beta = 0$, and in Theorem 2.1, [31] for arbitrary $\alpha > 0$ and $\beta = 0$. For arbitrary $\beta \neq 0$, the estimate is still valid. The proof is extremely similar to that of Theorem 2.1, [31], except that we interpret the functions in the space $H_{r,\alpha}$. For the sake of completeness, we include the proof here.

Theorem 2.1. *For any $u \in H_{r,\alpha}(\mathbb{R})$ and any integer $0 \leq \mu \leq r$, we have*

$$|u - P_N u|_\mu \lesssim \alpha^{\mu-r} N^{\frac{\mu-r}{2}} ||u||_{r,\alpha}, \tag{2.11}$$

where $|u|_\mu := ||\partial_x^\mu u||$ are the semi-norms.

Proof. By induction, we first show that for $\mu = 0$. For any integer $r \geq 0$,

$$||u - P_N u||^2 = \sum_{n=N+1}^{\infty} \hat{u}_n^2 = \sum_{n=N+1}^{\infty} \lambda_{n+1}^{-r} \lambda_{n+1}^{r} \hat{u}_n^2 \lesssim \alpha^{-2r} N^{-r} ||u||_{r,\alpha}^2. \tag{2.12}$$

Suppose for $1 \leq \mu \leq r$, (2.11) holds for $\mu - 1$. We need to show that (2.11) is also valid for μ. It is clear that

$$|u - P_N u|_\mu \leq |\partial_x u - P_N \partial_x u|_{\mu-1} + |P_N \partial_x u - \partial_x P_N u|_{\mu-1}. \tag{2.13}$$

On the one hand, due to the assumption for $\mu - 1$, we apply (2.11) to $\partial_x u$ and replace μ and r with $\mu - 1$ and $r - 1$, respectively:

$$|\partial_x u - P_N \partial_x u|_{\mu-1} \leq \alpha^{\mu-r} N^{\frac{\mu-r}{2}} ||\partial_x u||_{r-1,\alpha} \lesssim \alpha^{\mu-r} N^{\frac{\mu-r}{2}} ||u||_{r,\alpha}, \tag{2.14}$$

where the last inequality holds with the observation that

$$||\partial_x u||_{r-1,\alpha}^2 = \sum_{n=0}^{\infty} \lambda_{n+1}^{r-1} \widehat{(\partial_x u)}_n^2$$

and

$$\widehat{(\partial_x u)}_n = \int_{\mathbb{R}} \partial_x u H_n^{\alpha,\beta} dx = -\int_{\mathbb{R}} u \partial_x H_n^{\alpha,\beta} dx \overset{(2.5)}{=} \frac{\sqrt{\lambda_{n+1}}}{2} \int_{\mathbb{R}} u H_{n+1}^{\alpha,\beta} dx - \frac{\sqrt{\lambda_n}}{2} \int_{\mathbb{R}} u H_{n-1}^{\alpha,\beta} dx,$$

$$= \frac{\sqrt{\lambda_{n+1}}}{2} \hat{u}_{n+1} - \frac{\sqrt{\lambda_n}}{2} \hat{u}_{n-1}.$$

On the other hand, we have

$$\begin{aligned} P_N \partial_x u - \partial_x P_N u &= P_N \left(\sum_{n=0}^{\infty} \hat{u}_n \partial_x H_n^{\alpha,\beta} \right) - \sum_{n=0}^{N} \hat{u}_n \partial_x H_n^{\alpha,\beta} \\ &\overset{(2.5)}{=} -\frac{1}{2} \sum_{n=0}^{N-1} \sqrt{\lambda_{n+1}} \hat{u}_n H_{n+1}^{\alpha,\beta} + \frac{1}{2} \sum_{n=0}^{N+1} \sqrt{\lambda_n} \hat{u}_n H_{n-1}^{\alpha,\beta} - \left[-\frac{1}{2} \sum_{n=0}^{N} \sqrt{\lambda_{n+1}} \hat{u}_n H_{n+1}^{\alpha,\beta} + \frac{1}{2} \sum_{n=0}^{N} \sqrt{\lambda_n} \hat{u}_n H_{n-1}^{\alpha,\beta} \right] \\ &= \frac{1}{2} \sqrt{\lambda_{N+1}} \left[\hat{u}_N H_{N+1}^{\alpha,\beta} + \hat{u}_{N+1} H_N^{\alpha,\beta} \right]. \end{aligned}$$

This yields that

$$|P_N \partial_x u - \partial_x P_N u|_{\mu-1}^2 \lesssim \lambda_{N+1} \left(\hat{u}_N^2 \left| H_{N+1}^{\alpha,\beta} \right|_{\mu-1}^2 + \hat{u}_{N+1}^2 \left| H_N^{\alpha,\beta} \right|_{\mu-1}^2 \right), \tag{2.15}$$

due to the property of semi-norms. Moreover, we estimate $\hat{u}_k^2$ and $|H_k^{\alpha,\beta}|_{\mu-1}^2$, for $k = N, N+1$:

$$\hat{u}_N^2 \leq \sum_{n=N}^{\infty} \hat{u}_n^2 \leq ||u - P_{N-1}u||^2 \overset{(2.12)}{\lesssim} \alpha^{-2r}N^{-r}||u||_{r,\alpha}^2. \tag{2.16}$$

Similarly, $\hat{u}_{N+1}^2 \lesssim \alpha^{-2r}N^{-r}||u||_r^2$. And

$$\left|H_N^{\alpha,\beta}\right|_{\mu-1}^2 = \left|\left|\partial_x^{\mu-1}H_N^{\alpha,\beta}(x)\right|\right|^2 \overset{(2.10)}{\lesssim} \left|\left|H_N^{\alpha,\beta}(x)\right|\right|_{\mu-1,\alpha}^2 = \lambda_{N+1}^{\mu-1}, \tag{2.17}$$

since $\widehat{(H_N^{\alpha,\beta})}_k = \delta_{kN}$, for $k \in \mathbb{Z}^+$. Similarly, $|H_{N+1}^{\alpha,\beta}|_{\mu-1}^2 \lesssim \alpha^{-1}\lambda_{N+2}^{\mu-1}$. Substituting (2.16) and (2.17) into (2.15), we get

$$|P_N\partial_x u - \partial_x P_N u|_{\mu-1}^2 \lesssim \alpha^{2\mu-2r}N^{\mu-r}||u||_{r,\alpha}^2, \tag{2.18}$$

by the fact that $\lambda_N = 2N\alpha^2$. Combining (2.13), (2.14) and (2.18), we arrive at the conclusion. $\square$

3. Time-dependent Galerkin Hermite spectral method (TGHSM) to nonlinear convection-diffusion equations

Let us consider the following nonlinear convection–diffusion equation:

$$\begin{cases} u_t + a_1 g(u)u_x - a_2 u_{xx} + a_3 u_{xxx} = f(x,t), & \text{for } (x,t) \in \mathbb{R} \times [0,T] \\ u(x,0) = u_0(x), \end{cases} \tag{3.1}$$

where $a_2 \geq 0$, a_1, a_3 are arbitrary real parameters, and $g(\cdot) \in C(\mathbb{R})$ has the primitive function $G(\cdot)$.

The weak formulation of (3.1) is

$$\begin{cases} \langle \partial_t u, \varphi \rangle - a_1 \langle G(u), \partial_x \varphi \rangle + a_2 \langle \partial_x u, \partial_x \varphi \rangle - a_3 \langle \partial_{xx} u, \partial_x \varphi \rangle = \langle f, \varphi \rangle, & \text{for } (x,t) \in \mathbb{R} \times [0,T] \\ u(x,0) = u_0(x), \end{cases} \tag{3.2}$$

The THGSM of solving (3.1) is to find $u_N \in \mathcal{R}_N(t)$ such that

$$\begin{cases} \langle \partial_t u_N, \varphi \rangle - a_1 \langle G(u_N), \partial_x \varphi \rangle + a_2 \langle \partial_x u_N, \partial_x \varphi \rangle - a_3 \langle \partial_{xx} u_N, \partial_x \varphi \rangle = \langle f, \varphi \rangle, & \text{for } (x,t) \in \mathbb{R} \times [0,T] \\ u_N(x,0) = P_N^{\alpha(0),\beta(0)}u_0(x) := u_{N,0}, \end{cases} \tag{3.3}$$

for any $\varphi \in \mathcal{R}_N(t)$.

We shall investigate the stability and convergence analysis under the following assumption:

Assumption 1. $G(\cdot)$ has the primitive function $\tilde{G}(\cdot)$, and $\tilde{G}(0) = 0$.

Assumption 2. $|g(x)| \lesssim 1 + |x|^s$, and $|G(x)| \lesssim 1 + |x|^{s+1}$, $\forall x \in \mathbb{R}$, for some $s \geq 1$.

These two assumptions are not limited in the sense that they are satisfied by many important physical model equations, for example, the heat equation, the Burgers' equation and the Korteweg–de Vries Burger's (KdVB) equation. The first two equations are well-known and have been studied widely for a long time; while the last one is the model equation derived by Su and Gardner [27], and first studied by [24], which describes a wide class of nonlinear systems in the weak nonlinearity and long wavelength approximations, since it contains both damping and dispersion. This model equation also has been used in the study of wave propagation through a liquid-filled elastic tube [17] and for a description of shallow water waves on a viscous fluid [18]. The existence and uniqueness of the global smooth solution of KdVB equation have been established in [16].

3.1. A priori estimates

Lemma 3.4. If $u_0 \in L^2(\mathbb{R})$ and $f \in L^2\left([0,T]; L^2(\mathbb{R})\right)$, and *Assumption 1* holds, then

$$||u_N||^2(t) + 2a_2 e^T \int_0^t ||\partial_x u_N||^2(s)ds \leq e^T \left(||u_0||^2 + \int_0^t ||f||^2(s)ds\right). \tag{3.4}$$

Proof. Take the test function $\varphi = 2u_N$ in (3.3), we have

$$2\langle \partial_t u_N, u_N \rangle - 2a_1 \langle G(u_N), \partial_x u_N \rangle + 2a_2 \langle \partial_x u_N, \partial_x u_N \rangle - 2a_3 \langle \partial_{xx} u_N, \partial_x u_N \rangle = 2\langle f, u_N \rangle.$$

Under Assumption 1, we have

$$\langle G(u_N), \partial_x u_N \rangle = \tilde{G}(u_N)|_{-\infty}^{\infty} = 0,$$

due to the fact that $\lim_{x\to\pm\infty} u_N(x,t) = 0$, for all $t \in [0,T]$. And we also obtain that

$$\langle \partial_{xx} u_N, \partial_x u_N \rangle = \frac{1}{2}\int_{\mathbb{R}} \partial_x[(\partial_x u_N)^2]dx = 0, \tag{3.5}$$

since $u_N \in \mathcal{R}_N$, i.e., $\lim_{x \to \pm\infty} \partial_x^r u_N(x, t) = 0$, for any $r \in \mathbb{N}$ and $t \in [0, T]$. That is,

$$\frac{d}{dt} ||u_N||^2 + 2a_2 ||\partial_x u_N||^2 = 2 \int_{\mathbb{R}} f u_N dx \leq ||f||^2 + ||u_N||^2,$$

The Gronwall's inequality yields that

$$||u_N||^2 (t) + 2a_2 \int_0^t e^{t-s} ||\partial_x u_N||^2 (s) ds \leq e^t ||u_N||^2 (0) + \int_0^t e^{t-s} ||f||^2 (s) ds,$$

which implies (3.4). □

3.2. Stability

We shall consider the stability of (3.3) in the sense of Guo [12], since it is impossible to prove the stability of (3.3) in the sense of Courant et al. [8]. Let us assume that f and $u_{N,0}$ have the errors $\tilde{f}$ and $\tilde{u}_{N,0}$, respectively. They will introduce error in u_N, denoted as $\tilde{u}_N$. The error satisfies the following equation:

$$\begin{cases} \langle \partial_t \tilde{u}_N, \varphi \rangle - a_1 \langle G(u_N) - G(u_N + \tilde{u}_N), \partial_x \varphi \rangle + a_2 \langle \partial_x \tilde{u}_N, \partial_x \varphi \rangle - a_3 \langle \partial_{xx} \tilde{u}_N, \partial_x \varphi \rangle = \langle \tilde{f}, \varphi \rangle \\ \tilde{u}_N(x, 0) = \tilde{u}_{N,0}, \end{cases} \tag{3.6}$$

where $(x, t) \in \mathbb{R} \times [0, T]$, for all $\varphi \in \mathcal{R}_N$. By taking $\varphi = 2\tilde{u}_N$ in (3.6), it yields that

$$\frac{d}{dt} ||\tilde{u}_N||^2 - 2a_1 \langle g(u_N + \theta\tilde{u}_N)\tilde{u}_N, \partial_x \tilde{u}_N \rangle + 2a_2 ||\partial_x \tilde{u}_N||^2 = 2\langle \tilde{f}, \tilde{u}_N \rangle \leq ||\tilde{f}||^2 + ||\tilde{u}_N||^2, \tag{3.7}$$

due to the similar argument as in (3.5). The second term on the left-hand side of (3.7) can be easily estimated as

$$\begin{aligned} |\langle g(u_N + \theta\tilde{u}_N)\tilde{u}_N, \partial_x \tilde{u}_N \rangle| &\lesssim \frac{1}{\epsilon} ||g(u_N + \theta\tilde{u}_N)\tilde{u}_N||^2 + \epsilon ||\partial_x \tilde{u}_N||^2 \\ &\lesssim \frac{1}{\epsilon} \left(1 + ||u_N||_\infty^s + ||\tilde{u}_N||_\infty^s\right) ||\tilde{u}_N||^2 + \epsilon ||\partial_x \tilde{u}_N||^2 \\ &\lesssim C(\epsilon, s, ||u_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}, ||\tilde{u}_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}) ||\tilde{u}_N||^2 + \epsilon ||\partial_x \tilde{u}_N||^2, \end{aligned} \tag{3.8}$$

where

$$C(\epsilon, s, ||u_N||_{L^\infty(0,T;L^\infty(\mathbb{R})),||\tilde{u}_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}}) = \frac{1}{\epsilon} \left(||u_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}^s + ||\tilde{u}_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}^s \right).$$

This constant is finite, due to the fact that $u_N, \tilde{u}_N \in \mathcal{R}_N$, i.e., $||u_N||_\infty(t), ||\tilde{u}_N||_\infty(t) < \infty$, for any $t \in [0, T]$. Therefore, we arrive the following stability result:

Theorem 3.2. *Let u_N be the solution of (3.3) and $\tilde{u}_N$ be error induced by the error of the source term $\tilde{f}$ and that of the initial condition $\tilde{u}_{N,0}$. Then for any $t \leq T$, we have*

$$||\tilde{u}_N||^2 + C(a_1, a_2, \epsilon) \int_0^t ||\partial_x \tilde{u}_N(s)||^2 ds \lesssim \rho\left(\tilde{u}_{N,0}, \tilde{f}, t\right) + C(a_1, s, ||u_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}, ||\tilde{u}_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}, \epsilon) \int_0^t ||\tilde{u}_N(s)||^2 ds,$$

where

$$\begin{aligned} C(a_1, a_2, \epsilon) &= 2(a_2 - a_1\epsilon) > 0, \quad C(a_1, s, ||u_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}, ||\tilde{u}_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}, \epsilon) \\ &= \frac{2a_1}{\epsilon} \left(||u_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}^s + ||\tilde{u}_N||_{L^\infty(0,T;L^\infty(\mathbb{R}))}^s \right) + 1, \end{aligned}$$

and

$$\rho(\tilde{u}_{N,0}, \tilde{f}, t) = ||\tilde{u}_{N,0}||^2 + \int_0^t ||\tilde{f}(s)||^2 ds.$$

In particular, for all $t \leq T$,

$$||\tilde{u}_N||^2 + \int_0^t ||\partial_x \tilde{u}_N(s)||^2 ds \leq \rho(\tilde{u}_{N,0}, \tilde{f}, t) e^{C(a_1, u_N, T, \epsilon)t}.$$

This theorem implies that the error of the numerical solution is controlled by the errors of the initial data $\tilde{u}_{N,0}$ and the source term $\tilde{f}$. It means that (3.3) is of generalized stability in the sense of Guo [12].

3.3. Convergence analysis

Theorem 3.3. *If $u_0 \in H_r(\mathbb{R})$, and Assumptions 1–2 are satisfied, then for any $u \in L^\infty(0, T; H_{r,\alpha(t)}(\mathbb{R})) \cap L^2(0, T; H_{r,\alpha(t)}(\mathbb{R}))$ with $r > \max\{2, \frac{s}{4}\}$ and $N \gg 1$, we have*

1. *If $s < 8$, then*

$$||u - u_N||^2(t) \lesssim \alpha^{-2r}N^{-r}||u||^2_{L^\infty(0,T;H_{r,\alpha(t)}(\mathbb{R}))} + CN^{2-r}\int_0^t e^{(t-\tau)}||u(\tau)||^2_{r,\alpha(\tau)}d\tau; \tag{3.9}$$

2. *If $s \geq 8$, then*

$$||u - u_N||^2(t) \lesssim \alpha^{-2r}N^{-r}||u||^2_{L^\infty(0,T;H_{r,\alpha(t)}(\mathbb{R}))} + N^{\frac{s}{4}-r}\int_0^t e^{(t-\tau)}\alpha^{\frac{s}{2}-2r}(\tau)||u(\tau)||^2_{r,\alpha(\tau)}d\tau \tag{3.10}$$

where s is the growth rate of $g(\cdot)$ in Assumption 2 and C in (3.9) may depend on $\alpha, \alpha', s, r, ||u_N||_{L^\infty(0,T;L^2(\mathbb{R}))}, ||u||_{L^\infty(0,T;H_{r,\alpha(t)}(\mathbb{R}))}$ and $||u||_{L^2(0,T;H_{r,\alpha(t)}(\mathbb{R}))}$, etc.

Proof. Let $U_N = P_N^{\alpha,\beta}u$ for simpler notation. By the weak formulation (3.2) of the nonlinear convection-diffusion equation (3.1), U_N satisfies

$$\begin{cases} \langle \partial_t U_N, \varphi \rangle - a_1 \langle G(U_N), \partial_x\varphi \rangle + a_2 \langle \partial_x U_N, \partial_x\varphi \rangle - a_3 \langle \partial_{xx}U_N, \partial_x\varphi \rangle \\ \quad = -\langle \partial_t(u - U_N), \varphi \rangle + a_1 \langle G(u) - G(U_N), \partial_x\varphi \rangle \\ \qquad - a_2 \langle \partial_x(u - U_N), \partial_x\varphi \rangle + a_3 \langle \partial_{xx}(u - U_N), \partial_x\varphi \rangle + \langle f, \varphi \rangle \\ U_N(x, t) = P_N^{\alpha(0),\beta(0)}u_0(x). \end{cases} \tag{3.11}$$

Let $\eta_N = u_N - U_N$. It satisfies the following equation, which is the difference of (3.3) and (3.11):

$$\begin{cases} \langle \partial_t \eta_N, \varphi \rangle + a_2 \langle \partial_x \eta_N, \partial_x\varphi \rangle - a_3 \langle \partial_{xx}\eta_N, \partial_x\varphi \rangle \\ \quad = \langle \partial_t(u - U_N), \varphi \rangle + a_1 \langle G(u_N), \partial_x\varphi \rangle - a_1 \langle G(u) - G(U_N), \partial_x\varphi \rangle \\ \qquad + a_2 \langle \partial_x(u - U_N), \partial_x\varphi \rangle - a_3 \langle \partial_{xx}(u - U_N), \partial_x\varphi \rangle := \sum_{i=1}^5 G_i(\varphi) \\ \eta_N(x, t) = 0, \end{cases} \tag{3.12}$$

for any $\varphi(\cdot, t) \in \mathcal{R}_N(t)$. Let us take the test function $\varphi = 2\eta_N$, then

$$\text{L.H.S. of (3.12)} = 2\langle \partial_t \eta_N, \eta_N \rangle + 2a_2 \langle \partial_x \eta_N, \partial_x \eta_N \rangle - 2a_3 \langle \partial_{xx}\eta_N, \partial_x \eta_N \rangle \overset{(3.5)}{=} \frac{d}{dt}||\eta_N||^2 + 2a_2 ||\partial_x \eta_N||^2 \tag{3.13}$$

Next, we estimate the right-hand side of (3.12) term by term:

$$\frac{1}{2}G_1(2\eta_N) = \langle \partial_t(u - U_N), \eta_N \rangle \overset{(2.7)}{=} \left\langle \partial_t \left[\hat{u}_{N+1}H_{N+1}^{\alpha,\beta} + \hat{u}_{N+2}H_{N+2}^{\alpha,\beta}\right], \eta_N \right\rangle$$

$$= \left\langle \hat{u}_{N+1}\partial_t H_{N+1}^{\alpha,\beta}, \eta_N \right\rangle + \left\langle \partial_t\hat{u}_{N+1}H_{N+1}^{\alpha,\beta}, \eta_N \right\rangle + \left\langle \hat{u}_{N+2}\partial_t H_{N+2}^{\alpha,\beta}, \eta_N \right\rangle + \left\langle \partial_t\hat{u}_{N+2}H_{N+2}^{\alpha,\beta}, \eta_N \right\rangle$$

$$= \left\langle \hat{u}_{N+1}\partial_t H_{N+1}^{\alpha,\beta}, \eta_N \right\rangle + \left\langle \hat{u}_{N+2}\partial_t H_{N+2}^{\alpha,\beta}, \eta_N \right\rangle,$$

where the second and the last equalities are due to the fact that $\eta_N \in \mathcal{R}_N(t)$. Hence,

$$|G_1(2\eta_N)| \leq 2|\hat{u}_{N+1}| \cdot \left\|\partial_t H_{N+1}^{\alpha,\beta}(x, t)\right\| \cdot ||\eta_N|| + 2|\hat{u}_{N+2}| \cdot \left\|\partial_t H_{N+2}^{\alpha,\beta}(x, t)\right\| \cdot ||\eta_N|| \tag{3.14}$$

$$\overset{(2.7)}{\lesssim} \frac{|\alpha'|}{\alpha}N(|\hat{u}_{N+1}| + |\hat{u}_{N+2}|)||\eta_N||.$$

Note the fact that

$$|\hat{u}_{N+1}| \leq \left(\sum_{k=N+1}^\infty \hat{u}_k^2\right)^{\frac{1}{2}} = \left(\sum_{k=N+1}^\infty \lambda_{k+1}^{-r}\lambda_{k+1}^r\hat{u}_k^2\right)^{\frac{1}{2}} \leq \lambda_{N+2}^{-\frac{r}{2}}\left(\sum_{k=N+1}^\infty \lambda_{k+1}^r\hat{u}_k^2\right)^{\frac{1}{2}} \lesssim \alpha^{-r}(t)N^{-\frac{r}{2}}||u||_{r,\alpha(t)}.$$

Similarly, we have $|\hat{u}_{N+2}| \lesssim \alpha^{-r}(t)N^{-\frac{r}{2}}||u||_{r,\alpha(t)}$. We obtain the estimate of $|G_1(2\eta_N)|$:

$$|G_1(2\eta_N)| \lesssim |\alpha'(t)|\alpha^{-(r+1)}(t)N^{1-\frac{r}{2}}||u||_{r,\alpha(t)} \cdot ||\eta_N|| \lesssim ||\eta_N||^2 + |\alpha'(t)|^2\alpha^{-2-2r}(t)N^{2-r}||u||^2_{r,\alpha(t)}. \tag{3.15}$$

The second term $G_2(2\eta_N)$ on the right hand side of (3.12) can be estimated similarly as in (3.8):

$$|G_2(2\eta_N)| = 2a_1|\langle G(u_N), \partial_x\eta_N \rangle| \lesssim \frac{a_2}{2}||\partial_x\eta_N||^2 + C||u_N||^s_{L^\infty(0,T;L^2(\mathbb{R}))} + C, \tag{3.16}$$

due to Assumption 2. The third term $G_3(2\eta_N)$ on the right hand side of (3.12) yields:

$$|G_3(2\eta_N)| \lesssim |\langle G(u) - G(U_N), \partial_x \eta_N \rangle| = |\langle g(U_N + \theta(u - U_N))(u - U_N), \partial_x \eta_N \rangle|$$
$$\lesssim C(1 + ||u||_\infty^s + ||U_N||_\infty^s)||u - U_N||^2 + \frac{a_2}{2}||\partial_x \eta_N||^2$$

By the interpolation inequality (cf. [9]), we have

$$||U_N||_\infty \lesssim ||U_N||^{\frac{1}{2}}||\partial_x U_N||^{\frac{1}{2}} \lesssim \alpha^{\frac{1}{2}} N^{\frac{1}{4}}||U_N|| \leq \alpha^{\frac{1}{2}} N^{\frac{1}{4}}||u||,$$

by Lemma 2.2, where $|| \cdot ||_\infty$ is the L^∞ norm with respect to x. With this fact, we continue to estimate $G_3(2\eta_N)$:

$$|G_3(2\eta_N)| \lesssim \frac{a_2}{2}||\partial_x \eta_N||^2 + C\alpha^{-2r}(t)N^{-r}\left(||u||_\infty^s + \alpha^{\frac{s}{2}}(t)N^{\frac{s}{4}}||u||^s\right)||u||_{r,\alpha(t)}^2$$
$$\leq \frac{a_2}{2}||\partial_x \eta_N||^2 + C\alpha^{\frac{s}{2}-2r}(t)N^{\frac{s}{4}-r}||u||_{r,\alpha(t)}^2. \tag{3.17}$$

Other than the estimate of $||U_N||_\infty$, by Theorem 2.1, we have

$$||U_N||_\infty \lesssim ||U_N||^{\frac{1}{2}}||\partial_x U_N||^{\frac{1}{2}} \lesssim \alpha^{\frac{1-r}{2}}(t)\left(1 + N^{\frac{1-r}{2}}\right)^{\frac{1}{2}}||u||^{\frac{1}{2}}||u||_{r,\alpha(t)}^{\frac{1}{2}} \lesssim \alpha^{\frac{1-r}{2}}(t)||u||_{r,\alpha(t)}^{\frac{1}{2}}. \tag{3.18}$$

The rest two terms $G_i(2\eta_N)$, $i = 4, 5$, on the right-hand side of (3.12) can be easily bounded as follows:

$$|G_4(2\eta_N)| \lesssim \frac{a_2}{2}||\partial_x \eta_N||^2 + C||\partial_x(u - U_N)||^2 \leq \frac{a_2}{2}||\partial_x \eta_N||^2 + C\alpha^{2-2r}(t)N^{1-r}||u||_{r,\alpha(t)}^2; \tag{3.19}$$

and

$$|G_5(2\eta_N)| \lesssim \frac{a_2}{2}||\partial_x \eta_N||^2 + C||\partial_{xx}(u - U_N)||^2 \lesssim \frac{a_2}{2}||\partial_x \eta_N||^2 + C\alpha^{4-2r}(t)N^{2-r}||u||_{r,\alpha(t)}^2. \tag{3.20}$$

Substituting (3.15)–(3.17), (3.19) and (3.20) into the right-hand side of (3.12) and combining (3.13), we obtain that

$$\frac{d}{dt}||\eta_N||^2 \lesssim ||\eta_N||^2 + \left(CN^{2-r} + \alpha^{\frac{s}{2}-2r}(t)N^{\frac{s}{4}-r}\right)||u||_{r,\alpha(t)}^2,$$

where C is the function of t dependent of $\alpha, \alpha', s, r, ||u_N||_{L^\infty(0,T;L^2(\mathbb{R}))}$, etc. The results (3.9) and (3.10) in Theorem 3.3 are obtained by the Gronwall's inequality and Theorem 2.1. $\square$

4. Numerical results

In this section, we shall use our THGSM (3.2) to numerically solve three benchmark equations with $g(u) = \frac{u}{2}$ in (3.1), including the heat equation, the Burgers' equation and the KdVB equation with different coefficients a_i, $i = 1, 2, 3$ in (3.1). To illustrate the high accuracy of our method, we shall compare our numerical results of the first two equations with the ones obtained by using the similarity transformation technique [11] and by using the stabilized Hermite spectral method developed in [22].

In our scheme (3.3), we choose the test function $\varphi = H_m^{\alpha,\beta}(x,t) \in \mathcal{R}_N(t)$, for $m = 0, 1, \ldots, N$. It yields a system of ordinary differential equations (ODE) for $\hat{u}_m(t)$, $m = 0, 1, \ldots, N$. The second term on the left-hand side and the one on the right-hand side of (3.3) need to be computed numerically; while all the other terms can be explicitly written as sparse matrices, which are followed from the properties of the GHF. To be more precise, the first term gives

$$\frac{d}{dt}\vec{\hat{u}} + A_1\vec{\hat{u}}$$

where $\vec{\hat{u}} = [\hat{u}_0, \hat{u}_1, \ldots, \hat{u}_N]^T$, the superscript T means the transpose, where the matrix A_1 is explicitly given by (2.7) as:

$$A_1(i,j) = \begin{cases} -\frac{\alpha'}{\alpha}d(i)d(i-1), & \text{if } i = j+2 \\ \alpha\beta'd(i), & \text{if } i = j+1 \\ -\alpha\beta'd(j), & \text{if } i = j-1, \\ \frac{\alpha'}{\alpha}d(j)d(j-1), & \text{if } i = j-2 \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

$d(n) = \sqrt{\frac{n}{2}}$ as before. The third term on the left-hand side of (3.3) gives $a_2A_3\vec{\hat{u}}$, where

$$A_3(i,j) = \begin{cases} \alpha^2(t)[d^2(i+1) + d^2(i)], & \text{if } i = j \\ -\alpha^2(t)d(l+1)d(l+2), & l = \min\{i,j\}, \text{ if } |i-j| = 2, \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

Table 1
Errors of the heat equation at $T = 1$ with $N = 20$ using different methods.

Time step dt	Funaro and Kavian's scheme [11]	Ma et al.'s scheme [22] with $(\delta_0, \delta) = (1, 1)$	HSM scheme [21,31] with $\alpha = \frac{\sqrt{2}}{2}$	Our scheme (3.3) with $\alpha(t) = \frac{1}{\sqrt{2(t+1)}}$
250^{-1}	2.478E−03	2.958E−04	9.0032E−07	2.4045E−07
1000^{-1}	6.203E−04	1.189E−06	7.6286E−07	4.8534E−08
4000^{-1}	1.550E−04	1.177E−06	7.6213E−07	4.6247E−08
$16,000^{-1}$	3.886E−05	1.177E−06	7.6212E−07	4.6238E−08

by (2.6). And the last term on the left-hand side of (3.3) yields $-a_3 A_4 \vec{\hat{u}}$, where

$$
A_4(i, j) = \begin{cases}
\alpha^2(t)d(j+1)d(j+2)d(j+3), & \text{if } i = j+3 \\
-\alpha^2(t)[d^3(j+1) + d^2(j)d(j+1) + d(j+1)d^2(j+2)], & \text{if } i = j+1 \\
\alpha^2(t)[d^3(i+1) + d^2(i)d(i+1) + d(i+1)d^2(i+2)], & \text{if } i = j-1 \\
-\alpha^2(t)d(i+1)d(i+2)d(i+3), & \text{if } i = j-3 \\
0, & \text{otherwise}
\end{cases}
\tag{4.3}
$$

At last, the THGSM (3.3) yields the ODE system for the Fourier–Hermite coefficients $\vec{\hat{u}}$:

$$
\frac{d\vec{\hat{u}}}{dt} + A\vec{\hat{u}} + B(\vec{\hat{u}}) = \vec{\hat{f}},
\tag{4.4}
$$

where $B(\vec{\hat{u}})$ is given by the second term $\frac{a_1}{2}\langle \partial_x(u^2), \varphi \rangle$ of the left-hand side of (3.3), $\vec{\hat{f}}$ is the Fourier–Hermite coefficients of the source term $f(x, t)$ and $A = A_1 + A_3 + A_4$, A_1, A_3 and A_4 are given in (4.1)–(4.3), respectively. $\vec{\hat{f}}$ and $B(\vec{\hat{u}})$ need to be computed numerically. The equation (4.4) will be solved by using the Crank–Nicolson/Euler forward scheme in all the following examples. In particular, the Crank–Nicolson scheme is applied to all the linear terms, while the Euler forward scheme is applied to the nonlinear term $B(\vec{\hat{u}})$.

In all the numerical experiments below, we shall use the L^2-norm and the relative L^∞-norm to measure the accuracy. In our context, let us denote the L^2 error and the relative L^∞ error as

$$
E_N(t) = ||u_N(t) - u_{\text{exact}}(t)||
$$

and

$$
E_{N,\infty}(t) = \frac{\max_{0 \leq j \leq N} |u_N(y_j, t) - u_{\text{exact}}(y_j, t)|}{\max_{0 \leq j \leq N} |u_{\text{exact}}(y_j, t)|},
$$

respectively, where $\{y_j\}_{0 \leq j \leq N}$ are the Hermite–Gauss points. For comparison purpose, the examples in Section 4.1 and 4.2 are taken from [11,15,22]. The definitions of $E_N(t)$ and $E_{N,\infty}(t)$ are slightly different in various schemes. For example, the errors of scheme in [11] are in the weighted L^2-norm after the transformation, see definition on p. 615, [11]; those of scheme in [22] are in the weighted L^2-norm, see details on p. 71, [22].

4.1. Heat equation with source term

We consider (3.1) with $a_1 = a_3 = 0$, $a_2 = 1$, $g(u) = \frac{u}{2}$ and the following source term:

$$
f(x, t) = [x \cos x + (t+1) \sin x](t+1)^{-\frac{3}{2}} e^{-\frac{x^2}{4(t+1)}},
$$

which yields the heat equation. This equation with the initial condition $u(x, 0) = \sin(x)e^{-\frac{x^2}{4}}$ has been investigated in [11]. And its exact solution is

$$
u_{\text{exact}}(x, t) = \frac{\sin x}{\sqrt{t+1}} e^{-\frac{x^2}{4(t+1)}}.
$$

It is easy to see from the exact solution that it concentrates near the origin, so the translating factor β can be chosen to be zero. Moreover, the choice of the scaling factor α is extremely important for improving the accuracy of our scheme (3.3). We refer the readers to [21] for the guidelines to pick appropriate scaling factor. In this example, we shall investigate our scheme with the time-invariant scaling factor $\alpha = \frac{\sqrt{2}}{2}$ and the time-dependent one $\alpha(t) = \frac{1}{\sqrt{2(t+1)}}$. The HSM with time-invariant scaling factor is essentially developed in [21,31] and can be viewed as a special case of our THGSM. The accuracy of the numerical results obtained by our THGSM and the time-invariant scheme are compared with some previous results, including the ones in [11,22].

The results are list in Table 1. It is clear to see that our scheme is superior to the schemes in [11,22] in accuracy. And it is as we expected that the time-dependent scaling factor yields better resolution than the constant one, i.e. the scheme in

Table 2
Error of the heat equation by using the proposed scheme (3.3) with different time steps dt and truncation mode N.

Time step dt	N	$E_N(1)$	$E_{N,\infty}(1)$	Order
1E−1	40	1.7439E−04	1.2773E−04	
1E−2		1.7473E−06	1.2757E−06	$dt^{2.0005}$
1E−3		1.7473E−08	1.2757E−08	$dt^{2.0002}$
1E−4		1.7478E−10	1.2759E−10	$dt^{2.0001}$
1E−4	8	7.4E−03	9E−03	
	16	4.2446E−06	4.7275E−06	$N^{-10.77}$
	32	1.6540E−10	1.3012E−10	$N^{-12.71}$

Table 3
Errors of viscous Burger's equation at $T = e − 1$ with time step $dt = T \times 10^{-3}$. The error of scheme in [15] and that in [22] are $E_N(T)$ in their contexts, see details on p. 869, [15] and on p. 71, [22], respectively. The one in our scheme (3.3) is $E_N(T)$.

N	Guo and Xu's result [15]	Ma, Sun and Tang's scheme [22]	Our scheme (3.3)
8	1.381E−06	1.563E−05	1.7669E−06
16	1.381E−06	6.337E−07	1.1516E−07
32	1.381E−06	1.031E−07	1.3400E−07

[21,31]. From the viewpoint of computational complexity, our scheme can be implemented straightforward without any variable transformation. The matrix A in (4.4) is tri-diagonal, symmetric, and diagonally dominant. And the inverse of matrix A can be performed accurately and effectively inversed (part of the computation in Crank–Nicolson scheme).

To show the rate of convergence of (3.3), we list in Table 2 the numerical errors at $T = 1$ with various time steps dt and the truncation modes N. It confirms the theoretical prediction that the scheme (3.3) is of second-order accuracy in time and spectral accuracy in space.

4.2. Viscous Burger's equation

Besides the linear problems, our scheme is also effective in nonlinear problems. Let us take the viscous Burgers' equation as an example. The inviscid Burgers' equation suffers from the Gibbs' phenomenon, due to the lack of the diffusion term, i.e. $a_2 = 0$. The vanishing viscosity method combined with the spectral method, and the post-processing of the numerical solution are expected, see discussion in [29,30] and reference therein. The viscous Burgers' equation (3.1) with $a_1 = a_2 = 1$, $a_3 = 0$ and $g(u) = \frac{u}{2}$, has been studied in [15] by the variable transformation technique

$$y = \frac{x}{2\sqrt{t + 1}}, \quad s = \ln(t + 1)$$

that a soliton-like solution

$$u_{\text{exact}}(x, t) = e^{-\frac{x^2}{4(1+t)}} \operatorname{sech}^2\left(\frac{ax}{2(1 + t)} - b\ln(1 + t) - c\right)$$

is obtained with $a = 0.3$, $b = 0.5$ and $c = −3$, where the source term is

$$f(x, t) = -e^{-\frac{x^2}{2(1+t)}} \frac{\operatorname{sech}^4(\xi)}{1 + t}\left[\frac{x}{2} + a\tanh\xi\right]$$
$$+ e^{-\frac{x^2}{4(1+t)}} \frac{\operatorname{sech}^2(\xi)}{1 + t}\left[\frac{1 + t + a^2}{2(1 + t)} + 2b\tanh(\xi) - \tanh^2(\xi)\frac{3a^2}{2(1 + t)}\right],$$

$\xi = \frac{ax}{2(1+t)} - b\ln(1 + t) - c$.

In this example, the second term on the left-hand side of (3.3) is computed numerically. By taking $\varphi = H_m^{\alpha,\beta}$, $m = 0, \ldots, N$, the key component of this term

$$\left\langle \partial_x\left(u_N^2\right), \varphi \right\rangle \stackrel{(2.5)}{=} -\left\langle u_N^2, \partial_x H_m^{\alpha,\beta}\right\rangle = -\alpha d(m + 1)\left\langle u_N^2, H_{m+1}^{\alpha,\beta}\right\rangle + \alpha d(m)\left\langle u_N^2, H_{m-1}^{\alpha,\beta}\right\rangle$$

follows immediately with the numerical integral

$$\int_{\mathbb{R}} H_l^{\alpha,\beta} H_n^{\alpha,\beta} H_m^{\alpha,\beta},$$

$l, n, m = 0, \ldots, N$. We shall use the Gauss–Hermite quadrature rule to perform this integration.

In Table 3, we compare our time-dependent scheme with Guo–Xu's scheme in [15] and the time-dependent scheme in [22]. In this numerical experiment, we choose the same time step $dt = (e − 1) \times 10^{-3}$ and the total experimental time $T = e − 1$ as in

Table 4
Errors of viscous Burger's equation by using scheme in [22] and our scheme (3.3) with different time steps. The errors $E_N(t)$ and $E_{N,\infty}(t)$ in the scheme [22] and our scheme are defined slightly different.

Time step dt	Ma et al.'s scheme [22]			Time-dependent HSM (3.3)		
	$E_{40}(1)$	$E_{40,\infty}(1)$	Order	$E_{20}(1)$	$E_{20,\infty}(1)$	Order
1E−1	5.101E−04	4.677E−03		4.8044E−06	1.4264E−04	
1E−2	4.508E−06	4.548E−05	$dt^{2.05}$	4.1512E−07	1.0680E−05	$dt^{1.06}$
1E−3	4.454E−08	4.530E−07	$dt^{2.01}$	4.1065E−08	1.0324E−06	$dt^{1.03}$
1E−4	4.467E−10	4.372E−09	$dt^{2.00}$	4.1771E−09	9.7699E−08	$dt^{1.02}$

Table 5
Errors of viscous Burger's equation by using our scheme (3.3) with different truncation modes N.

Time step dt	N	$E_N(1)$	$E_{N,\infty}(1)$	Order
1E−5	5	2.4254E−06	9.4692E−05	
	15	1.1291E−08	4.5015E−07	$N^{-4.88}$
	25	4.3637E−10	1.0527E−08	$N^{-5.28}$

Table 6
Errors of KdVB equation with different time steps.

Time step dt	$E_{40}(1)$	$E_{40,\infty}(1)$	Order
1E−2	3.11E−02	1.6E−02	
1E−3	4.4076E−04	2.83E−04	$dt^{1.85}$
1E−4	3.4747E−04	2.9070E−04	$dt^{0.98}$

[15,22] for comparison. It is clear to see that our scheme is at least as good as the one in [22], and is more accurate than the one in [15].

Moreover, in Table 4 we compare our scheme with the one in [22] with different time steps dt. We choose the total experimental time $T = 1$, and the truncation modes $N = 40$ and 20, for the scheme in [22] and our THGSM, respectively. It is shown that they are almost as the same accuracy. In this table, we have also verified the first-order accuracy in time, due to the Euler forward scheme used in the nonlinear term $B(\vec{u})$. In Table 5, we illustrate the spectral accuracy of our scheme in space as we showed in Theorem 3.3.

4.3. KdVB equation

In this subsection, we shall reveal the effect of the time-dependent translating factor, which has been concealed in the theoretical proof of the convergence analysis. Let us consider the 1D KdVB equation (3.1) with $a_1 = a_2 = 1$, $a_3 = -\frac{1}{16}$, $g(u) = \frac{u}{2}$ and the source term

$$f(x, t) = -8\text{sech}^2\xi (2 - 3\text{sech}^2\xi + 2\tanh\xi \, \text{sech}^2\xi),$$

where $\xi = 2(x + t)$. It is easy to verify that

$$u_{\text{exact}}(x, t) = \text{sech}^2(\xi)$$

is the exact solution to the above KdVB equation. The Crank–Nicolson/Euler forward scheme is used as in the case of the viscous Burgers' equation.

In Table 6, we list the errors of numerical solutions to the KdVB equation with various time steps. The total experimental time is $T = 1$, and the truncation mode is $N = 40$. The scaling factor is chosen to be a constant $\alpha = 2\sqrt{2}$ and the translating factor is $\beta = 0$. It is shown that we can achieve almost the first-order accuracy in time. However, the approximate solution is less accurate than the one we obtained in the viscous Burgers' equation. The possible reason for this may be the drifting to the left of the exact solution.

To tackle this problem, we move our GHF accordingly by setting a time-dependent translating factor $\beta = -t$. We compare the results with those obtained without translating factor, i.e. $\beta = 0$. In Table 7, we display the errors of the scheme (3.3) with and without translating factor. The total experimental time is $T = 1$, and the scaling factor is chosen to be a constant $\alpha = 2\sqrt{2}$. Various truncation modes $N = 10, 20, \ldots, 50$ are numerically experimented. It is shown in Table 7 that the time-dependent translating factor yields better resolution. The spectral accuracy is verified in the scheme (3.3) for the KdVB equation in both with/without translating factor cases.

Table 7
Errors of the KdVB equation by using the proposed scheme (3.3) without translating factor and with time-dependent translating factor $\beta = -t$. Different truncation modes N are numerically experimented.

Time step dt	N	No translating factor $\beta = 0$			Time-dependent translating factor $\beta = -t$		
		$E_N(1)$	$E_{N,\infty}(1)$	Order	$E_N(1)$	$E_{N,\infty}(1)$	Order
1E−4	10	9.35E−02	7.54E−02		5.7E−03	2.6E−03	
	20	9.6E−03	8.2E−03	$N^{-3.28}$	2.8944E−04	1.2570E−04	
	30	1.6E−03	1.4E−03	$N^{-3.66}$	6.8894E−05	2.5007E−05	$N^{-4.05}$
	40	3.4747E−04	2.9070E−04	$N^{-4.17}$	1.3225E−05	4.5626E−06	$N^{-4.27}$
	50	9.0630E−05	7.1222E−05	$N^{-4.36}$	3.5135E−06	1.1749E−06	$N^{-4.49}$

5. Conclusions

In this paper, we investigate a time-dependent Hermite–Galerkin spectral method (THGSM) to solve the nonlinear convection–diffusion equations in the whole line. Many important physical model equations are within the framework in this paper. Our THGSM is formulated under the basis of the generalized Hermite functions (GHF), which includes the time-dependent scaling factor and translating factor in the definition. The advantages of this formulation are at least two folds: on the one hand, the proof of the stability and convergence analysis is much easier than the previous Hermite spectral method, say in [11,22], where they analyze in the weighted L^2 space. On the other hand, the implementation of our method is straightforward. No variable transformation technique is required. Furthermore, the derived ODE system in the linear equation is with extremely low computational cost, since the stiffness matrix is symmetric. The numerical experiments are carried out in some benchmark examples, including the heat equation, the viscous Burgers' equation and the KdVB equation. It is clear to see that our method surpasses nearly all the existing methods in accuracy.

Acknowledgments

References

[1] N. Adžić, Modified hermite polynomials in the spectral approximation for boundary layer problems, Bull. Aust. Math. Soc. 45 (1992) 267–276.
[2] W. Bao, S. Jin, P. Markowich, Numerical study of time-splitting spectral discretizations of nonlinear shrödinger equations in the semiclassical regimes, SIAM J. Sci. Comput. 25 (1) (2003) 27–64.
[3] W. Bao, J. Shen, A fourth-order time-splitting laguerre-hermite pseudo-spectral method for bose-einstein condensates, SIAM J. Sci. Comput. 26 (2005) 2010–2028.
[4] J. Boyd, Asymptotic coefficients of hermite function series, J. Comput. Phys. 54 (1984) 382–410.
[5] H. Brunner, Q. Hu, Optimal superconvergence results for delay integro-differential equations of pantograph type, SIAM J. Numer. Anal. 45 (2007) 986–1004.
[6] Q. Chen, D. Gottlieb, J. Hesthaven, Spectral methods based on prolate spheroidal wave functions for hyperbolic pdes, SIAM J. Numer. Anal. 43 (5) (2005) 1912–1933.
[7] N. Condette, C. Melcher, E. Süli, Spectral approximation of pattern-forming nonlinear evolution equations with double-well potentials of quadratic growth, Math. Comput. 80 (273) (2011) 205–233.
[8] R. Courant, K.O. Friedrichs, H. Levy, Über die partiellen differezengleichungen der mathematischen physik, Math. Annal. 100 (1928) 32–74.
[9] A. Friedman, Partial Differential Equations, Holt, Reihart and Winston, New York, 1969.
[10] J. Fok, B. Guo, T. Tang, Combined hermite spectral-finite difference method for the fokker-planck equation, Math. Comput. 71 (240) (2001) 1497–1528.
[11] D. Funaro, O. Kavian, Approximation of some diffusion evolution equation in unbounded domains by hermite function, Math. Comput. 37 (1991) 597–619.
[12] B.-Y. Guo, Generalized stability of discretization and its applications to numerical solutions of nonlnear differential equations, Computational Mathematics in China, Contemporary Mathematics vol. 163 (1994) 33–54.
[13] B.-Y. Guo, J. Shen, Laguerre-galerkin method for nonlinear partial differential equations on a semi-infinite interval, Numer. Math. 86 (2000) 635–654.
[14] B.-Y. Guo, J. Shen, C.-L. Xu, Spectral and pseudospectral approximation using hermite function: Application in dirac equation, Adv. Comput. Math. 19 (2003) 35–55.
[15] B.-Y. Guo, C.-L. Xu, Hermite pseudospectral method for nonlinear partial differential equations, ESAIM Math. Model. Numer. Anal. 34 (2000) 869–872.
[16] B. Guo, The global solutions for a class of generalized KDV equations, Acta Math. Sin. 25 (6) (1982) 641–656.
[17] R.S. Jonson, A nonlinear equation incorporating damping and dispersion, J. Phys. Mech. 42 (1970) 49–60.
[18] R.S. Jonson, Shallow water waves in a viscous fluid-the undular bore, Phys. Fluids 15 (1972) 1693–1699.
[19] I.K. Khabibrakhmanov, D. Summers, The use of generalized laguerre polynomials in spectral methods for nonlinear differential equations, Comput. Math. Appl. 36 (1998) 65–70.
[20] J. Lopez, J. Shen, An efficient spectral-projection method for the navier-stokes equations in cylindrical geometries i. axisymmetric cases, J. Comput. Phys. 139 (1998) 308–326.
[21] X. Luo, S.S.-T. Yau, Hermite spectral method to 1d forward kolmogorov equation and its application to nonlinear filtering problems, IEEE Trans. Automat. Control 58 (10) (2013) 2495–2507.
[22] H. Ma, W. Sun, T. Tang, Hermite spectral methods with a time-dependent scaling for parabolic equations in unbounded domains, SIAM J. Numer. Anal. 43 (1) (2005) 58–75.
[23] H. Ma, T. Zhao, A stabilized hermite spectral method for second-order differential equations in unbounded domains, Numer. Methods Partial Differential Equ. 23 (5) (2007) 968–983.

[24] J.C. Saut, Sur quelques generalizations de l'equation de korteweg-de vries, J. Math. Pures Appl. 58 (1979) 21–61.
[25] J.W. Schumer, J.P. Holloway, Vlasov simulations using velocity-scaled hermite representations, J. Comput. Phys 144 (1998) 626–661.
[26] J. Shen, T. Tang, L.-L. Wang, Spectral methods: Algorithm, analysis and applications, springer series in computational mathematics, vol. 41, Springer (2011).
[27] C.H. Su, C.S. Gardner, Kdv equation and generalization. iv. derivation of the korteweg-de vries and burgers' equations, J. Math. Phys. 10 (1969) 536–539.
[28] T. Tang, The hermite spectral method for gaussian-type functions, SIAM J. Sci. Comput. 14 (1993) 594–606.
[29] E. Tadmor, Convergence of spectral methods for nonlinear conservation laws, SIAM J. Numer. Anal. 26 (1989) 30–44.
[30] E. Tadmor, Shock capturing by the spectral viscosity method, Comput. Methods Appl. Mech. Eng. 80 (1990) 197–208.
[31] X.-M. Xiang, Z.-Q. Wang, Generalized hermite spectral method and its applications to problems in unbounded domains, SIAM J. Numer. Anal. 48 (4) (2010) 1231–1253.
[32] X.-M. Xiang, Z.-Q. Wang, Generalized hermite approximations and spectral method for partial differential equations in multiple dimensions, J. Sci. Comput. 57 (2) (2013) 229–253.