



An adaptive sparse grid local discontinuous Galerkin method for Hamilton-Jacobi equations in high dimensions

Wei Guo^{a,1}, Juntao Huang^{b,*}, Zhanjing Tao^{c,2}, Yingda Cheng^{d,3}

^a Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, 79409, USA

^b Department of Mathematics, Michigan State University, East Lansing, MI 48824, USA

^c School of Mathematics, Jilin University, Changchun, Jilin 130012, China

^d Department of Mathematics, Department of Computational Mathematics, Science and Engineering, Michigan State University, East Lansing, MI 48824, USA

ARTICLE INFO

Article history:

Available online 26 March 2021

Keywords:

Sparse grid
Adaptivity
Local discontinuous Galerkin
Hamilton-Jacobi equations
High dimensions

ABSTRACT

The Hamilton-Jacobi (HJ) equations arise in optimal control and many other applications. Oftentimes, such equations are posed in high dimensions, and this presents great numerical challenges. In this paper, we propose an adaptive sparse grid (also called adaptive multiresolution) local discontinuous Galerkin (DG) method for solving Hamilton-Jacobi equations in high dimensions. By using the sparse grid techniques, we can treat moderately high dimensional cases. Adaptivity is incorporated to capture kinks and other local structures of the solutions. Two classes of multiwavelets including the orthonormal Alpert's multiwavelets and the interpolatory multiwavelets are used to achieve multiresolution. Numerical tests in up to four dimensions are provided to validate the performance of the method.

© 2021 Elsevier Inc. All rights reserved.

1. Introduction

In this paper, we consider the Hamilton-Jacobi (HJ) equation

$$\phi_t + H(\nabla\phi) = 0, \quad (1.1)$$

on the bounded domain $[0, 1]^d$ in arbitrary d dimension, subject to initial condition $\phi(\mathbf{x}, 0) = \phi_0(\mathbf{x})$ and appropriate boundary conditions. The HJ equation has diverse applications in science and engineering, such as optimal control, seismic waves, crystal growth, robotic navigation, image processing, calculus of variations, among others [35]. In this paper, we develop a class of adaptive sparse grid (also called adaptive multiresolution) discontinuous Galerkin (DG) methods for approximating the viscosity solution of (1.1). The concept of viscosity solution was developed by Crandall and Lions in [12,11] to single out the physically relevant weak solution. Under certain assumptions, the viscosity solution can be interpreted by the Hopf formula [18], and the numerical approximation to the viscosity solution is of practical interest [42]. It is well known that the viscosity solution of the HJ equation is only Lipschitz continuous and may develop discontinuous derivatives in finite time

* Corresponding author.

E-mail addresses: weimath.guo@ttu.edu (W. Guo), huangj75@msu.edu (J. Huang), zjtao@jlu.edu.cn (Z. Tao), yicheng@msu.edu (Y. Cheng).

¹ Research is supported by NSF grant DMS-1830838.

² Research is supported by NSF Grant 12001231.

³ Research is supported by NSF grants DMS-1453661 and DMS-1720023.

regardless of smoothness of the initial condition. Various numerical methods for solving (1.1) have been developed in the literature [42], such as the monotone methods [13,1,32], the essentially non-oscillatory (ENO) methods [38,39], the weighted ENO (WENO) and the Hermite WENO (HWENO) methods [28,47,40] among many others. In this work, we choose to use the DG discretization due to their distinguished advantages in handling geometry, boundary conditions and accommodating adaptivity, which are highly desirable for efficiently solving the HJ equation. Several DG schemes have been proposed in the literature [24,33,5,34,46,22,6,30]. Here, we use the local DG (LDG) method developed by Yan and Osher [46], which has provable property for the piecewise constant case and is easy to implement, although we remark that the extensions to other DG formulations are also possible.

Beyond the need to capture the viscosity solution, another major numerical challenge for HJ equation is that it is often posed in high dimensions, and any standard numerical discretization becomes inefficient due to the curse of dimensionality. Recent years have seen a surge of interests in designing numerical solutions of HJ equations in high dimensions. Various approaches have been proposed, including those using sparse grid [3,19,29], model order reduction [31], tensor decomposition [17], Hopf formula [16,7,8] and machine learning [23,37,14,15], to name a few. Some of the work above is feasible for HJ equations in hundreds of dimensions for some special cases, and continued efforts to develop efficient numerical solvers for high-dimensional HJ PDEs constitutes a vibrant research area due to their wide applications in control and differential games.

In this paper, we take the sparse grid approach [4], which has been used in [3,19,29] for computations in moderately high dimension. The scheme we proposed relies on multiresolution analysis (MRA) [36] and is designed to be high order accurate. In a line of research, we have developed a family of adaptive sparse grid (or adaptive multiresolution) DG methods for linear transport equations with application to kinetic equations [21], hyperbolic conservation laws [25], and wave equations [26]. By incorporating MRA and the sparse grid ideas, our methods are able to efficiently capture small-scale structures, and more importantly, work very well in high dimensions. In particular, in [25,26], we use two classes of multiwavelets to achieve MRA. The tensor-product Alpert's multiwavelets are used as the DG function space, following the approach developed in [45,20,21] for linear equations. Besides, the interpolatory multiwavelets for MRA quadrature [44] are used for computations of nonlinear terms. Numerical experiments for benchmark tests in up to four dimension verify the efficiency and efficacy of the method in capturing the viscosity solution of the HJ equations.

The rest of the paper is organized as follows. In Section 2, we review the fundamentals of Alpert's and interpolatory multiwavelets. In Section 3, we present the LDG method for the HJ equation with MRA. Some theoretical results and implementation details are discussed. Section 4 contains numerical examples. In Section 5, we include the conclusion of this paper.

2. Multiresolution analysis and multiwavelets

In this section, we first review the fundamentals of MRA of DG approximation spaces and the associated multiwavelets. Two classes of multiwavelets, namely the L^2 orthonormal Alpert's multiwavelets [2] and the interpolatory multiwavelets [44], are considered. We also introduce a set of key notations used throughout the paper.

2.1. Alpert's multiwavelets

We start with the construction of Alpert's multiwavelets [2], which have been employed to develop a class of sparse grid DG methods for solving PDEs in high dimensions [45,20]. For a unit sized interval $\Omega = [0, 1]$, we define a set of nested grids $\Omega_0, \Omega_1, \dots$, for which the n -th level grid Ω_n consists of 2^n uniform cells

$$I_n^j = (2^{-n}j, 2^{-n}(j+1)], \quad j = 0, \dots, 2^n - 1.$$

Denote $I_{-1} = [0, 1]$. The piecewise polynomial space of degree at most k on grid Ω_n for $n \geq 0$ is denoted by

$$V_n^k := \{v : v \in P^k(I_n^j), \forall j = 0, \dots, 2^n - 1\}. \quad (2.1)$$

Observing the nested structure

$$V_0^k \subset V_1^k \subset V_2^k \subset V_3^k \subset \dots,$$

we can define the multiwavelet subspace W_n^k , $n = 1, 2, \dots$ as the orthogonal complement of V_{n-1}^k in V_n^k with respect to the L^2 inner product on $[0, 1]$, i.e.,

$$V_{n-1}^k \oplus W_n^k = V_n^k, \quad W_n^k \perp V_{n-1}^k.$$

By letting $W_0^k := V_0^k$, we obtain a hierarchical decomposition $V_n^k = \bigoplus_{0 \leq l \leq n} W_l^k$, i.e., MRA of space V_n^k . A set of orthonormal basis can be defined on W_l^k as follows. When $l = 0$, the basis $v_{i,0}^0(x)$, $i = 0, \dots, k$ are the normalized shifted Legendre polynomials in $[0, 1]$. When $l > 0$, the Alpert's orthonormal multiwavelets [2] are employed as the bases and denoted by

$$v_{i,l}^j(x), \quad i = 0, \dots, k, \quad j = 0, \dots, 2^{l-1} - 1.$$

We then follow a tensor-product approach to construct the hierarchical finite element space in multi-dimensional space. Denote $\mathbf{l} = (l_1, \dots, l_d) \in \mathbb{N}_0^d$ as the mesh level in a multivariate sense, where \mathbb{N}_0 denotes the set of nonnegative integers, we can define the tensor-product mesh grid $\Omega_{\mathbf{l}} = \Omega_{l_1} \otimes \dots \otimes \Omega_{l_d}$ and the corresponding mesh size $h_{\mathbf{l}} = (h_{l_1}, \dots, h_{l_d})$. Based on the grid $\Omega_{\mathbf{l}}$, we denote $I_{\mathbf{l}}^{\mathbf{j}} = \{\mathbf{x} : x_m \in (h_m j_m, h_m(j_m + 1)), m = 1, \dots, d\}$ as an elementary cell, and

$$\mathbf{V}_{\mathbf{l}}^k := \{\mathbf{v} : \mathbf{v} \in Q^k(I_{\mathbf{l}}^{\mathbf{j}}), \mathbf{0} \leq \mathbf{j} \leq 2^{\mathbf{l}} - \mathbf{1}\} = V_{l_1, x_1}^k \times \dots \times V_{l_d, x_d}^k$$

as the tensor-product piecewise polynomial space, where $Q^k(I_{\mathbf{l}}^{\mathbf{j}})$ represents the collection of polynomials of degree up to k in each dimension on cell $I_{\mathbf{l}}^{\mathbf{j}}$. If we use equal mesh refinement of size $h_N = 2^{-N}$ in each coordinate direction, the grid and space will be denoted by Ω_N and \mathbf{V}_N^k , respectively. Based on a tensor-product construction, the multi-dimensional increment space can be defined as

$$\mathbf{W}_{\mathbf{l}}^k = W_{l_1, x_1}^k \times \dots \times W_{l_d, x_d}^k.$$

The basis functions in multi-dimensions are defined as

$$v_{\mathbf{i}, \mathbf{l}}^{\mathbf{j}}(\mathbf{x}) := \prod_{m=1}^d v_{l_m, l_m}^{j_m}(x_m), \quad (2.2)$$

for $\mathbf{l} \in \mathbb{N}_0^d$, $\mathbf{j} \in B_{\mathbf{l}} := \{\mathbf{j} \in \mathbb{N}_0^d : \mathbf{0} \leq \mathbf{j} \leq \max(2^{\mathbf{l}-1} - \mathbf{1}, \mathbf{0})\}$ and $\mathbf{1} \leq \mathbf{i} \leq \mathbf{k} + \mathbf{1}$.

Using the notation of

$$\|\mathbf{l}\|_1 := \sum_{m=1}^d l_m, \quad \|\mathbf{l}\|_{\infty} := \max_{1 \leq m \leq d} l_m,$$

and the same component-wise arithmetic operations and relations as defined in [45], we reach the decomposition

$$\mathbf{V}_N^k = \bigoplus_{\substack{\|\mathbf{l}\|_{\infty} \leq N \\ \mathbf{l} \in \mathbb{N}_0^d}} \mathbf{W}_{\mathbf{l}}^k. \quad (2.3)$$

On the other hand, a standard choice of sparse grid space [45,20] is

$$\hat{\mathbf{V}}_N^k = \bigoplus_{\substack{\|\mathbf{l}\|_1 \leq N \\ \mathbf{l} \in \mathbb{N}_0^d}} \mathbf{W}_{\mathbf{l}}^k \subset \mathbf{V}_N^k. \quad (2.4)$$

We skip the details about the property of the space, but refer the readers to [45,20]. In Section 3, we will describe the adaptive scheme which adapts a subspace of \mathbf{V}_N^k according to the numerical solution, hence offering more flexibility and efficiency.

2.2. Interpolatory multiwavelets

Alpert's multiwavelets described in Section 2.1 are associated with the L^2 projection operator. The interpolatory multiwavelets introduced in [44] are constructed based on interpolation operators and also essential for efficient computation of integrals in the DG formulation, especially in high dimensions. In this work, only Lagrange interpolation is considered, while we note that Hermite interpolation can also be used but its implementation is more involved. The details are provided below.

We first define the set of $M + 1$ distinct interpolation points on the interval $I = [0, 1]$ at zeroth mesh level by $X_0 = \{x_i\}_{i=0}^M \subset I$. We defer the discussion of the relations between M and k to Section 3.2.

The interpolation points at mesh level $n \geq 1$, X_n can be obtained correspondingly as

$$X_n = \{x_{i,n}^j := 2^{-n}(x_i + j), \quad i = 0, \dots, M, \quad j = 0, \dots, 2^n - 1\}.$$

We require the points to be nested, i.e.

$$X_0 \subset X_1 \subset X_2 \subset X_3 \subset \dots \quad (2.5)$$

This can be achieved by requiring $X_0 \subset X_1$.

Given the interpolation points, we define the basis functions on the zeroth level grid as Lagrange interpolation polynomials of degree $\leq M$ which satisfy the property:

$$\phi_i(x_{i'}) = \delta_{i,i'},$$

for $i, i' = 0, \dots, M$. It is easy to see that $\text{span}\{\phi_i, i = 0, \dots, M\} = V_0^M$. With the basis function at mesh level zero, we can define the basis functions at mesh level $n \geq 1$:

$$\phi_{i,n}^j := \phi_i(2^n x - j), \quad i = 0, \dots, M, \quad j = 0, \dots, 2^n - 1,$$

which form a complete basis set for V_n^M .

We now introduce the hierarchical representations and the interpolatory multiwavelets. Define $\tilde{X}_0 := X_0$ and $\tilde{X}_n := X_n \setminus X_{n-1}$ for $n \geq 1$, then we have the decomposition

$$X_n = \tilde{X}_0 \cup \tilde{X}_1 \cup \dots \cup \tilde{X}_n.$$

Denote the points in \tilde{X}_1 by $\tilde{X}_1 = \{\tilde{x}_i\}_{i=0}^M$. Then the points in \tilde{X}_n for $n \geq 1$ can be represented by

$$\tilde{X}_n = \{\tilde{x}_{i,n}^j := 2^{-(n-1)}(\tilde{x}_i + j), \quad i = 0, \dots, M, \quad j = 0, \dots, 2^{n-1} - 1\}.$$

For notational convenience, we let $\tilde{W}_0^M := V_0^M$. The increment function space \tilde{W}_n^M for $n \geq 1$ is introduced as a function space that satisfies

$$V_n^M = V_{n-1}^M \oplus \tilde{W}_n^M, \quad (2.6)$$

and is defined through the multiwavelets $\psi_i \in V_1^M$ that satisfies

$$\psi_i(x_{i'}) = 0, \quad \psi_i(\tilde{x}_{i'}) = \delta_{i,i'},$$

for $i, i' = 0, \dots, M$. Then \tilde{W}_n^M is given by

$$\tilde{W}_n^M = \text{span}\{\psi_{i,n}^j := \psi_i(2^{n-1}x - j), \quad i = 0, \dots, M, \quad j = 0, \dots, 2^{n-1} - 1\}.$$

The multi-dimensional construction follows similar lines as in Section 2.1. We let

$$\tilde{\mathbf{W}}_1^M = \tilde{W}_{l_1, x_1}^M \times \dots \times \tilde{W}_{l_d, x_d}^M,$$

then

$$\mathbf{V}_N^M = \bigoplus_{\substack{\|\mathbf{l}\|_\infty \leq N \\ \mathbf{l} \in \mathbb{N}_0^d}} \tilde{\mathbf{W}}_1^M,$$

while the sparse grid approximation space is

$$\hat{\mathbf{V}}_N^M = \bigoplus_{\substack{\|\mathbf{l}\|_1 \leq N \\ \mathbf{l} \in \mathbb{N}_0^d}} \tilde{\mathbf{W}}_1^M.$$

Note that the constructions by Alpert's multiwavelets and the interpolatory multiwavelets deduce the same sparse grid space because of the same nested structure. Finally, the interpolation operator in multidimension is defined as $\mathcal{I}_N^M : C(\Omega) \rightarrow \mathbf{V}_N^M$:

$$\mathcal{I}_N^M[f](\mathbf{x}) = \sum_{\substack{\|\mathbf{n}\|_\infty \leq N \\ \mathbf{0} \leq \mathbf{j} \leq \max(2^{n-1}-1, \mathbf{0}) \\ \mathbf{0} \leq \mathbf{i} \leq \mathbf{M}}} b_{\mathbf{i}, \mathbf{n}}^{\mathbf{j}} \psi_{\mathbf{i}, \mathbf{n}}^{\mathbf{j}}(\mathbf{x}),$$

where the multi-dimensional basis functions $\psi_{\mathbf{i}, \mathbf{n}}^{\mathbf{j}}(\mathbf{x})$ are defined in the same approach as (2.2) by tensor products:

$$\psi_{\mathbf{i}, \mathbf{n}}^{\mathbf{j}}(\mathbf{x}) := \prod_{m=1}^d \psi_{i_m, n_m}^{j_m}(x_m). \quad (2.7)$$

The fast algorithms which transform point values at interpolation points to hierarchical coefficients $\{b_{\mathbf{i}, \mathbf{n}}^{\mathbf{j}}\}$ are given in [44].

For the sparse grid space $\hat{\mathbf{V}}_N^M$ or any adaptively chosen subspace of \mathbf{V}_N^M , the interpolation operator, which is denoted by \mathcal{I}_h^M in later sections, can be defined accordingly, by taking only multiwavelet basis functions that belong to that space. For the detailed formulas of the interpolation points and the associated interpolatory multiwavelets used in this work, we refer readers to [25,26].

3. Adaptive multiresolution LDG scheme

In this section, we present the adaptive multiresolution LDG method for simulating the HJ equation (1.1). We start with reviewing the LDG formulation by Yan and Osher in [46]. Then, by incorporating MRA and multiwavelets introduced in the previous section, we define our scheme.

3.1. LDG formulation

We consider periodic boundary conditions for simplicity, while the method can be adapted to other non-periodic boundary conditions. For illustrative purposes, we first introduce a set of shorthand notations. Denote by Γ the union of the boundaries for all the elements in the partition Ω_N . The jump and average of $q \in L^2(\Gamma)$ are defined as

$$[q] = q^- \mathbf{n}^- + q^+ \mathbf{n}^+, \quad \{q\} = \frac{1}{2}(q^- + q^+),$$

where \mathbf{n} is the unit normal. ‘ $-$ ’ and ‘ $+$ ’ represent that the directions of the vector point to interior and exterior at e , respectively. Note that $[q] \in [L^2(\Gamma)]^d$, and we let $[q]_m$ denote the m -th component of $[q]$.

The key idea in [46] is to employ the standard LDG methodology, see e.g. [10], to reconstruct the first derivatives of ϕ , i.e., ϕ_{x_m} , $m = 1, \dots, d$. In particular, the LDG method computes two piecewise polynomials p_m^1 and p_m^2 , both approximating ϕ_{x_m} but using opposite one-sided numerical fluxes; that is, given ϕ_h we seek p_m^τ , $m = 1, \dots, d$, $\tau = 1, 2$ in \mathbf{V} such that for all $w_h \in \mathbf{V}$

$$\int_{\Omega} p_m^\tau w_h \, d\mathbf{x} = - \int_{\Omega} \phi_h (w_h)_{x_m} \, d\mathbf{x} + \sum_{e \in \Gamma} \int_e \hat{\phi}_m^\tau [w_h]_m \, ds, \quad (3.1)$$

where the numerical fluxes are defined as

$$\hat{\phi}_m^1 = \{\phi_h\} + \frac{1}{2}[\phi_h]_m, \quad \hat{\phi}_m^2 = \{\phi_h\} - \frac{1}{2}[\phi_h]_m.$$

Note that p_m^1 and p_m^2 carry the information of ϕ_{x_m} from opposite directions. Hence, when the solution is smooth, p_m^1 and p_m^2 are almost identical, while if the solution involves nonsmooth corners, then p_m^1 and p_m^2 can be very different.

Then the semi-discrete scheme for solving (1.1) is defined as follows: seek $\phi_h \in \mathbf{V}$ such that, for all $v \in \mathbf{V}$,

$$\int_{\Omega} (\phi_h)_t v \, d\mathbf{x} + \int_{\Omega} \hat{H}(p_1^1, p_1^2, p_2^1, p_2^2, \dots, p_d^1, p_d^2) v \, d\mathbf{x} = 0, \quad (3.2)$$

where \hat{H} denotes a monotone numerical Hamiltonian that approximates H , and p_m^τ , $\tau = 1, 2$, $m = 1, \dots, d$ are given in (3.1). In the simulations, we employ the following global Lax-Friedrichs Hamiltonian

$$\hat{H}(p_1^1, p_1^2, p_2^1, p_2^2, \dots, p_d^1, p_d^2) = H(\bar{p}_1, \bar{p}_2, \dots, \bar{p}_d) - \sum_{m=1}^d \frac{\alpha_m}{2} (p_m^2 - p_m^1),$$

where $\bar{p}_m = \frac{1}{2}(p_m^1 + p_m^2)$ and

$$\alpha_m = \max_{q_1, \dots, q_d} \left| \frac{\partial H(q_1, \dots, q_d)}{\partial q_m} \right|$$

with the maximum being taken over the whole domain.

Depending on the choice of space \mathbf{V} , we obtain several LDG methods for (1.1) with distinct properties. If $\mathbf{V} = \mathbf{V}_N^k$, we recover the full grid LDG scheme in [46] on tensor-product meshes. If $\mathbf{V} = \hat{\mathbf{V}}_N^k$, then we obtain the sparse grid LDG method. If \mathbf{V} is chosen adaptively, we have the adaptive sparse grid scheme. Noteworthy, besides the LDG formulation, we can employ other DG formulations as well, such as the direct DG method [6] and the indirect DG methods [24,22]. The LDG formulation used is comparatively simpler to implement under the MRA framework.

If the Hamiltonian H is linear, then the HJ equation (1.1) degenerates to a transport equation with constant coefficients, and the formulation (3.2) together with (3.1) is nothing but a standard upwind DG scheme. The results established in [20] can be adapted directly to the linear HJ equation that if the solutions are adequately smooth in terms of the mixed norm, then the sparse grid DG method using space $\hat{\mathbf{V}}_n^k$ is convergent of order $k+1$ with a polylogarithmic factor.

3.2. Semi-discrete scheme with multiresolution interpolation

For nonlinear problems, one major difficulty of implementation of formulation (3.2) is to compute the volume integral efficiently and accurately, especially in high dimensions. Naive implementation of numerical quadratures is inefficient due to the hierarchical structure of multiwavelets. To address the challenge, we follow the idea in [41,25] and interpolate the numerical Hamiltonian \hat{H} by using the multiresolution Lagrange interpolation discussed in Section 2.2. In particular, we have the following modified formulation with interpolation. We find $\phi_h \in \mathbf{V}$ so that for all $v \in \mathbf{V}$

$$\int_{\Omega} (\phi_h)_t v \, d\mathbf{x} + \int_{\Omega} \mathcal{I}_h^M \left(\hat{H}(p_1^1, p_1^2, p_2^1, p_2^2, \dots, p_d^1, p_d^2) \right) v \, d\mathbf{x} = 0. \quad (3.3)$$

By doing so, not only can we apply the unidirectional principle to facilitate the computation, but also fast algorithms can be utilized to improve efficiency. We omit the details regarding the fast algorithm and refer readers to [25].

Note that the interpolation procedure in the scheme formulation plays a role as a high order MRA numerical quadrature. There exist two types of points, namely inner points and interface points [44]. It is observed that the schemes using the interface points are more stable than those using the inner points with the same order accuracy (see, e.g. [25]), and hence we choose to use the interface points in the simulations. We also remark if the Hamiltonian H is not smooth and $k > 1$, then to ensure stability, one must employ a very high order quadrature, i.e. large M , for accurate computation of the volume integral. This is ascribed to the fact that the large quadrature error due to nonsmoothness of H may pollute the numerical viscosity and lead to instability. This drawback is observed in [46], and the authors further coupled a nonlinear limiter to restore stability. In this paper, we propose to properly regularize the Hamiltonian so that the interpolation \mathcal{I}_h^M is adequately accurate for stability. The details will be presented in Section 4. Another possible approach is to add artificial viscosity, as done for solving conservation laws [25].

To preserve the accuracy of the original DG scheme, the interpolation operator $\mathcal{I}_h^M(\cdot)$ needs to reach certain accuracy. Following [9], we can write the DG scheme with interpolation (3.3) into the semi-discrete form as

$$\frac{d\phi_h}{dt} = L_h(\phi_h). \quad (3.4)$$

Here $L_h(\cdot)$ is an operator onto \mathbf{V} and is a discrete approximation of $-H(\nabla\phi)$ which satisfies

$$\int_{\Omega} L_h(\phi_h) v \, d\mathbf{x} + \int_{\Omega} \mathcal{I}_h^M \left(\hat{H}(p_1^1, p_1^2, p_2^1, p_2^2, \dots, p_d^1, p_d^2) \right) v \, d\mathbf{x} = 0, \quad (3.5)$$

for all $v \in \mathbf{V}$ with p_m^τ , $\tau = 1, 2, m = 1, \dots, d$ determined by (3.1). Using similar techniques as in [9,27], we have the following proposition on local truncation error of the sparse grid method with $\mathbf{V} = \hat{\mathbf{V}}_N^k$. The proof is omitted for brevity.

Proposition 3.1 (Local truncation error analysis). *If the interpolation operator \mathcal{I}_h^M in (3.3) has the accuracy of order $|\log_2 h_N|^d h_N^{k+1}$ for sufficiently smooth functions, then the local truncation error of the semi-discrete DG scheme with interpolation (3.3) is of order $|\log_2 h_N|^d h_N^{k+1}$. To be more precise, for sufficiently smooth Hamiltonian H and function ϕ , the sparse grid DG method with interpolation (3.3) has the truncation error:*

$$\|L_h(\phi) + H(\nabla\phi)\|_{L^2(\Omega)} \leq C |\log_2 h_N|^d h_N^{k+1}. \quad (3.6)$$

Here, we use C to denote a generic constant that may depend on the solution u , but does not depend on N .

The proposition indicates that, to preserve the order accuracy of the original scheme, we should use $M \geq k$. Hence, in the simulation we let $M \geq k$. Meanwhile, many Hamiltonians are non-smooth functions, and indeed we observe numerically that we need $M > k$ for those cases. This will be further discussed in Section 4. Note that if $M \geq k$, then \mathcal{I}_h^M preserves linear terms. Hence, there is no difference if we apply to \mathcal{I}_h^M only to the nonlinear part or to the entire numerical Hamiltonian.

For time discretization, we employ the third order strong-stability-preserving Runge-Kutta (RK) scheme [43] to advance the semi-discrete scheme (3.4). The adaptive procedure follows the technique developed in [3,21] to determine the space \mathbf{V} that dynamically evolves over time. The details are omitted for brevity. The main idea is that in light of the distinguished property of the orthonormal multiwavelets, we keep track of multiwavelet coefficients, i.e. L^2 norms of ϕ_h , as a natural error indicator for refining and coarsening, aiming to efficiently capture the viscosity solution of (1.1) which may develop discontinuous derivatives.

4. Numerical examples

In this section, we present a collection of numerical examples to demonstrate the performance of the proposed adaptive sparse grid LDG method for solving the HJ equation. We consider numerical examples up to $d = 4$ with smooth and nonsmooth Hamiltonian, and with smooth and nonsmooth viscosity solutions. Noteworthy, we may need to tune M for optimal

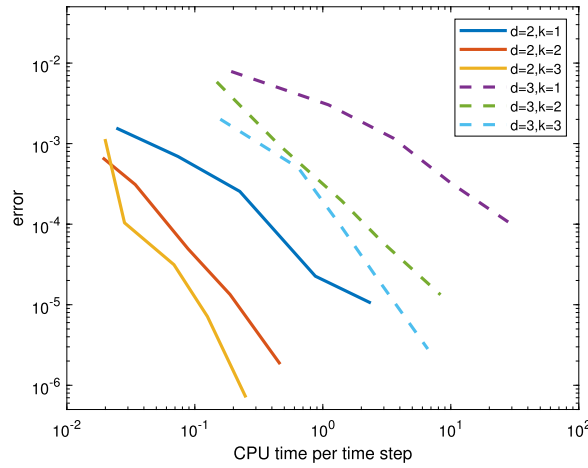


Fig. 4.1. Example 4.1, L_2 error versus CPU time per time step.

performance. In particular, we observe that for some numerical tests, we can simply take $M = k$ to achieve satisfactory results and maintain the original accuracy of the DG method, while for the some other tests, we may need to take larger M to ensure good performance. In all numerical simulations, the value of M is taken between k and $k + 2$.

Example 4.1. Consider the following Burgers' equation in d -dimension

$$\begin{cases} \phi_t + \frac{1}{2} \left(\sum_{m=1}^d \phi_{x_m} \right)^2 = 0, & \mathbf{x} \in [0, 1]^d, \\ \phi(\mathbf{x}, 0) = -\frac{1}{2\pi} \cos \left(2\pi \sum_{m=1}^d x_m \right), \end{cases} \quad (4.1)$$

with periodic boundary conditions.

At $T = 0.01$ for $d = 2$ and $T = 0.005$ for $d = 3$, the solutions are still smooth, and we summarize the convergence study of the sparse grid method with $\hat{\mathbf{V}}_N^k$ in Tables 4.1-4.2, including the L_2 errors and the associated orders of accuracy, with various configurations of k and M . It is observed that larger M leads to smaller error magnitude as expected. Slight order reduction is observed for $d = 2$, and it becomes more severe for $d = 3$. This is because, as time evolves, the viscosity solution of (4.1) develops larger and larger mixed derivatives especially in high dimensions. Hence, it may not be optimal to use the sparse grid space $\hat{\mathbf{V}}_N^k$ for approximating the viscosity solution. In Tables 4.3-4.4, we report the convergence study for the adaptive method for $d = 2, 3$, respectively. In particular, by fixing the maximum mesh level $N = 7$, two rates of convergence are calculated [3]. The first one is with respect to the error threshold:

$$R_{\epsilon_l} = \frac{\log(e_{l-1}/e_l)}{\log(\epsilon_{l-1}/\epsilon_l)},$$

and the second is with respect to DoF:

$$R_{\text{DoF}_l} = \frac{\log(e_{l-1}/e_l)}{\log(\text{DoF}_l/\text{DoF}_{l-1})}.$$

For the full grid counterpart, we have $R_{\text{DoF}} = (k + 1)/d$ for smooth solutions. It is observed that $R_{\epsilon} < 1$, which is similar to the Burgers' equation [25]. Furthermore, using larger k is beneficial, as the method with larger k requires less DoF to attain a certain level of accuracy. To further demonstrate the efficiency of the adaptive algorithm, we report L_2 errors versus the average CPU cost per time step for $k = 1, 2, 3$ and $d = 2, 3$ in Fig. 4.1. It is observed that, to achieve a desired level of accuracy, the adaptive method with a larger k requires less CPU time as expected, and moreover, the CPU cost is approximately proportional to the corresponding DoF count reported in Tables 4.3-4.4. In Fig. 4.2, we plot the solution profile together with active elements for the adaptive methods with configuration $N = 6$, $\epsilon = 10^{-6}$ at $T = 0.01$. The sparse grid elements of space $\hat{\mathbf{V}}_N^k$ are also plotted for comparison. It is observed that, when the solution is smooth, the performance of the adaptive method is similar to that of the sparse grid method.

At $T = 0.04$ for $d = 2$ and $T = 0.02$ for $d = 3$, the viscosity solutions have developed discontinuous derivatives. In Fig. 4.3, we plot the solution profiles computed by the sparse grid method and the adaptive method for $d = 2$. We set $k = 2$, $M = 2$

Table 4.1Example 4.1, $d = 2$. Sparse grid. $T = 0.01$.

	N	$M = 1$		$M = 2$		$M = 3$	
		L_2 error	order	L_2 error	order	L_2 error	order
$k = 1$	3	2.63E-02	–	1.99E-02	–	1.99E-02	–
	4	7.87E-03	1.74	5.88E-03	1.76	5.75E-03	1.79
	5	3.82E-03	1.04	2.42E-03	1.28	2.24E-03	1.36
	6	1.77E-03	1.11	8.27E-04	1.55	6.64E-04	1.76
	7	7.95E-04	1.16	3.47E-04	1.25	2.10E-04	1.66
	N	$M = 2$		$M = 3$		$M = 4$	
		L_2 error	order	L_2 error	order	L_2 error	order
$k = 2$	3	5.68E-03	–	2.84E-03	–	2.81E-03	–
	4	1.21E-03	2.23	3.75E-04	2.92	3.66E-04	2.94
	5	3.10E-04	1.97	1.42E-04	1.40	1.41E-04	1.37
	6	5.28E-05	2.56	1.97E-05	2.84	1.96E-05	2.85
	7	9.58E-06	2.46	4.36E-06	2.18	4.32E-06	2.18
	N	$M = 3$		$M = 4$		$M = 5$	
		L_2 error	order	L_2 error	order	L_2 error	order
$k = 3$	3	1.14E-03	–	6.28E-04	–	6.10E-04	–
	4	1.68E-04	2.76	6.86E-05	3.19	6.58E-05	3.21
	5	2.59E-05	2.70	1.45E-05	2.24	1.44E-05	2.20
	6	2.10E-06	3.63	7.94E-07	4.19	7.84E-07	4.20
	7	2.77E-07	2.92	1.50E-07	2.41	1.49E-07	2.39

Table 4.2Example 4.1, $d = 3$. Sparse grid. $T = 0.005$.

	N	$M = 1$		$M = 2$		$M = 3$	
		L_2 error	order	L_2 error	order	L_2 error	order
$k = 1$	3	4.67E-02	–	4.02E-02	–	3.78E-02	–
	4	4.19E-02	0.16	2.42E-02	0.73	2.29E-02	0.72
	5	2.49E-02	0.75	1.20E-02	1.02	8.51E-03	1.43
	6	1.10E-02	1.18	5.38E-03	1.15	3.87E-03	1.14
	7	9.64E-03	0.19	2.49E-03	1.11	1.92E-03	1.01
	8	4.44E-03	1.12	1.50E-03	0.74	6.28E-04	1.61
	N	$M = 2$		$M = 3$		$M = 4$	
		L_2 error	order	L_2 error	order	L_2 error	order
$k = 2$	3	3.34E-02	–	9.22E-03	–	7.46E-03	–
	4	1.71E-02	0.96	3.69E-03	1.32	3.24E-04	1.20
	5	6.93E-03	1.31	1.32E-03	1.49	1.21E-03	1.43
	6	2.07E-03	1.75	5.49E-04	1.26	5.28E-04	1.19
	7	8.07E-04	1.36	1.57E-04	1.80	1.54E-04	1.78
	8	1.91E-04	2.08	4.09E-05	1.94	3.98E-05	1.95
	N	$M = 3$		$M = 4$		$M = 5$	
		L_2 error	order	L_2 error	order	L_2 error	order
$k = 3$	3	8.86E-03	–	3.56E-03	–	2.48E-03	–
	4	2.97E-03	1.58	1.10E-03	1.70	8.62E-04	1.53
	5	9.97E-04	1.57	3.64E-04	1.59	2.93E-04	1.56
	6	3.08E-04	1.70	9.78E-05	1.90	8.57E-05	1.77
	7	6.49E-05	2.24	2.17E-05	2.17	2.04E-05	2.07
	8	1.44E-05	2.17	5.02E-06	2.12	4.85E-06	2.07

and $\epsilon = 10^{-5}$, and the maximum mesh level $N = 6$ for the adaptive method. It is observed the sparse grid method is able to capture the main structure of the solution, but severe oscillations appear due to lack of mesh resolution around the corners, while the adaptive method is able to effectively capture the viscosity solution by adding more DoF in the nonsmooth region. Hence, the sparse grid method with fixed space $\hat{\mathbf{V}}_N^k$ cannot reliably approximate the nonsmooth viscosity solution. Afterwards, we will only focus on the performance of adaptive method. For $d = 3$, we set $k = 2$, $M = 3$ and $\epsilon = 10^{-5}$, and the maximum mesh level $N = 6$, and plot the results generated by the adaptive method in Fig. 4.4, including the 2D cuts of the solution and the associated active elements at final time. Similar results to $d = 2$ are observed.

Table 4.3Example 4.1, $d = 2$. Adaptive sparse grid. $T = 0.01$. $M = k$.

	ϵ	DoF	L_2 error	R_ϵ	R_{DoF}
$k = 1$	1.00E-03	448	1.56E-03	–	–
	1.00E-04	1376	6.92E-04	0.35	0.73
	1.00E-05	3520	2.55E-04	0.43	1.06
	1.00E-06	10240	2.26E-05	1.05	2.27
	1.00E-07	18688	1.05E-05	0.33	1.28
$k = 2$	1.00E-03	270	6.67E-04	–	–
	1.00E-04	720	3.10E-04	0.33	0.78
	1.00E-05	1548	4.93E-05	0.80	2.40
	1.00E-06	3492	1.34E-05	0.57	1.60
	1.00E-07	7704	1.83E-06	0.86	2.51
$k = 3$	1.00E-03	192	1.14E-03	–	–
	1.00E-04	480	1.04E-04	1.04	2.62
	1.00E-05	896	3.14E-05	0.52	1.92
	1.00E-06	1856	7.14E-06	0.64	2.03
	1.00E-07	3136	7.07E-07	1.00	4.41

Table 4.4Example 4.1, $d = 3$. Adaptive sparse grid. $T = 0.005$. $M = k$.

	ϵ	DoF	L_2 error	R_ϵ	R_{DoF}
$k = 1$	1.00E-03	2432	7.87E-03	–	–
	1.00E-04	14864	3.03E-03	0.41	0.53
	1.00E-05	44656	1.17E-03	0.41	0.87
	1.00E-06	152176	3.25E-04	0.56	1.04
	1.00E-07	380976	9.05E-05	0.56	1.39
$k = 2$	1.00E-03	2646	5.84E-03	–	–
	1.00E-04	8208	9.84E-04	0.77	1.57
	1.00E-05	21816	1.96E-04	0.70	1.65
	1.00E-06	55404	6.11E-05	0.51	1.25
	1.00E-07	133569	1.33E-05	0.66	1.74
$k = 3$	1.00E-03	2048	2.01E-03	–	–
	1.00E-04	6400	5.01E-04	0.60	1.22
	1.00E-05	16384	9.26E-05	0.73	1.80
	1.00E-06	35584	2.30E-05	0.60	1.79
	1.00E-07	99840	2.82E-06	0.91	2.04

Example 4.2. Consider the following HJ equation with a nonconvex Hamiltonian

$$\begin{cases} \phi_t - \cos\left(\sum_{m=1}^d \phi_{x_m} + 1\right) = 0, & \mathbf{x} \in [0, 1]^d, \\ \phi(\mathbf{x}, 0) = -\frac{1}{2\pi} \cos\left(2\pi \sum_{m=1}^d x_m\right) \end{cases} \quad (4.2)$$

with periodic boundary conditions.

In Tables 4.5–4.6, we report the convergence rates for the adaptive method for $d = 2$ and $d = 3$ at $T = 0.01$ and $T = 0.005$, respectively. Similar results are observed to the previous example. In Fig. 4.5, we report the solution profile together with the active elements used at $T = 0.06$ when the viscosity solution has developed nonsmooth corners. In this simulations, we set $N = 6$ and $\epsilon = 10^{-5}$. Again, the adaptive method is able to efficiently and correctly capture the sharp corners. In Fig. 4.6, we plot the results for $d = 3$ at $T = 0.03$ with configuration parameters $k = 2$, $M = 3$, maximum level $N = 6$, and $\epsilon = 10^{-6}$. High resolution result is observed.

Example 4.3. We consider the following two-dimensional nonlinear problem

$$\begin{cases} \phi_t + \phi_{x_1} \phi_{x_2} = 0, & \mathbf{x} \in [0, 1]^2, \\ \phi(\mathbf{x}, 0) = -\frac{1}{2\pi} (\sin(2\pi x_1) + \cos(2\pi x_2)) \end{cases} \quad (4.3)$$

with periodic boundary conditions.

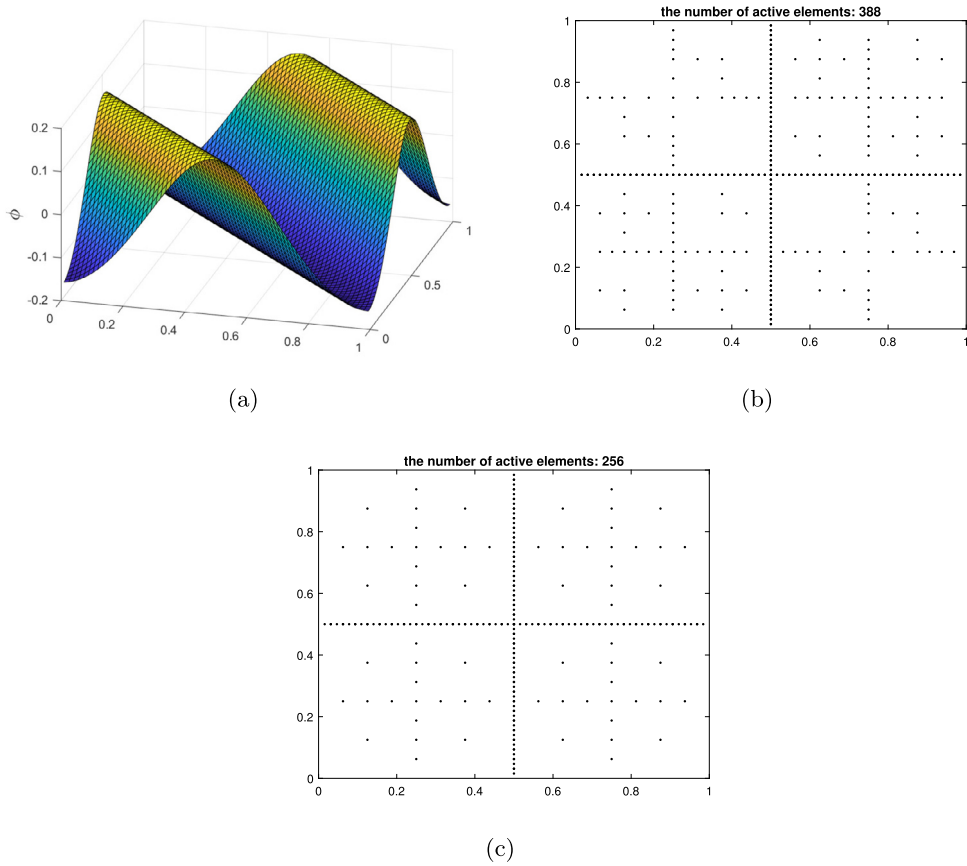


Fig. 4.2. Example 4.1, $d = 2$, $k = 2$, $M = 2$, $T = 0.01$, $N = 6$, $\varepsilon = 10^{-6}$. (a) Numerical solutions by adaptive sparse grid. (b) Active elements of the adaptive algorithm. (c) Sparse grid elements with $N = 6$.

Table 4.5

Example 4.2, $d = 2$. Adaptive sparse grid. $T = 0.01$. $M = k$.

	ϵ	DoF	L_2 error	R_ϵ	R_{DoF}
$k = 1$	1.00E-03	464	1.47E-03		
	1.00E-04	1616	4.60E-04	0.51	0.93
	1.00E-05	3840	1.66E-04	0.44	1.18
	1.00E-06	9056	2.37E-05	0.85	2.27
	1.00E-07	17440	7.86E-06	0.48	1.68
$k = 2$	1.00E-03	288	1.43E-03		
	1.00E-04	720	3.20E-04	0.65	1.64
	1.00E-05	1656	9.24E-05	0.54	1.49
	1.00E-06	3924	1.79E-05	0.71	1.90
	1.00E-07	8406	4.00E-06	0.65	1.97
$k = 3$	1.00E-03	192	1.53E-03		
	1.00E-04	512	1.85E-04	0.92	2.15
	1.00E-05	960	4.56E-05	0.61	2.23
	1.00E-06	2048	1.50E-05	0.48	1.47
	1.00E-07	3968	1.64E-06	0.96	3.34

Note that unlike the previous two examples, the problem is genuinely nonlinear, and the Hamiltonian is smooth but nonconvex. When $T = 0.03$, the solution is still smooth, and we are able to test the convergence for the adaptive method. In the simulation, we set maximum level $N = 6$, $k = 2$, $M = 3$. It is observed in Table 4.7 that the method is able to achieve very accurate results by using a few DoFs. The convergence performance is similar to the previous examples. In Fig. 4.7, we plot the solution at $T = 0.2$, when the viscosity solution becomes nonsmooth. It is observed that the adaptive method captures the corners correctly and efficiently, as compared with the results by other popular methods, see e.g. [34,6].

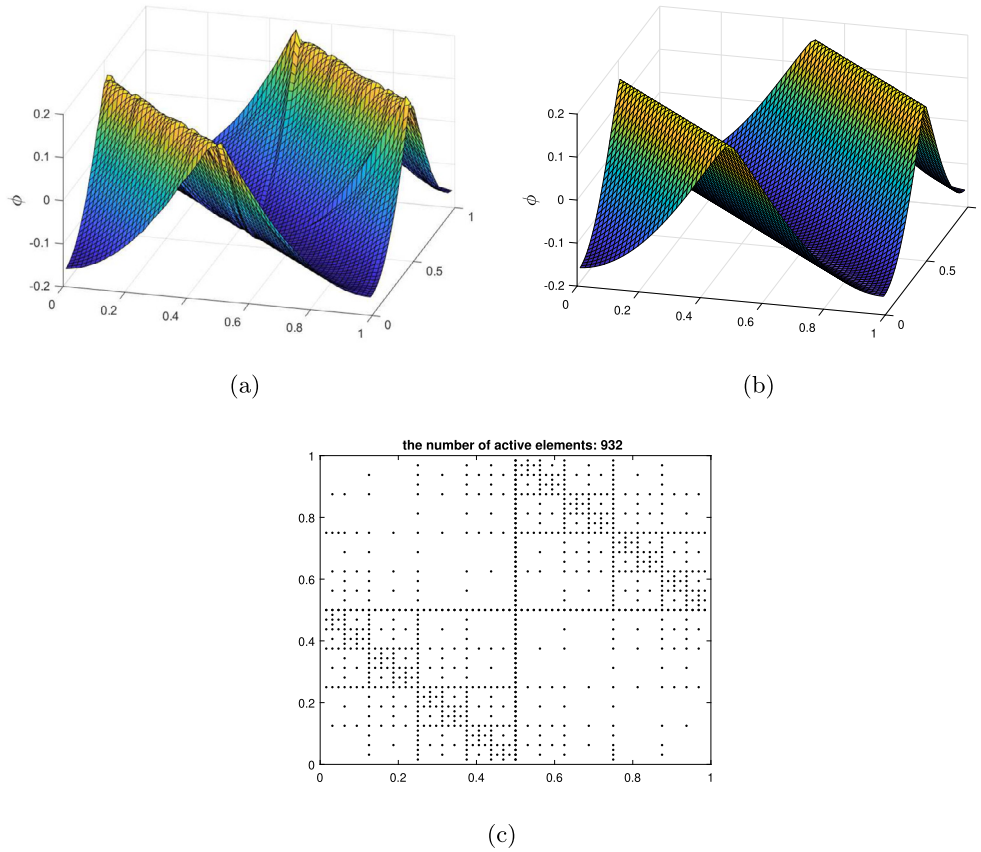


Fig. 4.3. Example 4.1, $d = 2$, $k = 2$, $M = 2$, $T = 0.04$, $N = 6$, $\varepsilon = 10^{-5}$. (a) Numerical solution by sparse grids. (b) Numerical solutions by adaptive sparse grid. (c) Active elements.

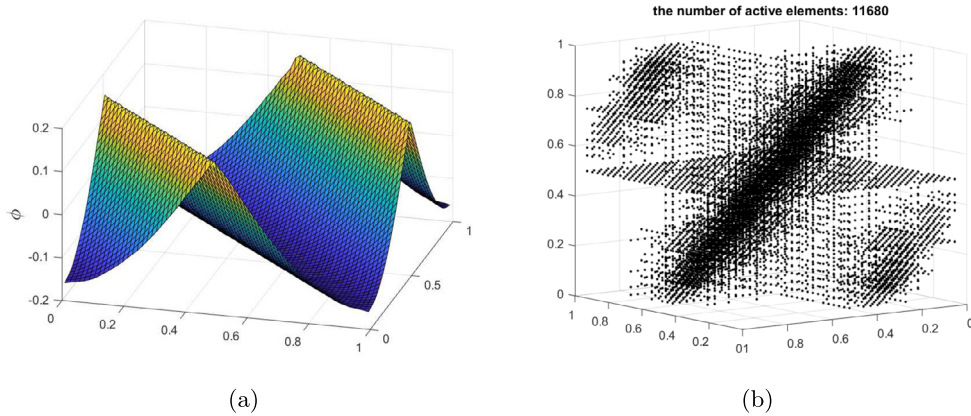


Fig. 4.4. Example 4.1, $d = 3$, $k = 2$, $M = 3$, $T = 0.02$, $N = 6$, $\varepsilon = 10^{-5}$. (a) 2D-cuts of the numerical solution at $x_3 = 0$. (b) Active elements.

Example 4.4. We consider the classic nonlinear Eikonal equation

$$\begin{cases} \phi_t + \|\nabla \phi\| = 0, & \mathbf{x} \in [0, 1]^d \\ \phi(\mathbf{x}, 0) = g(\|\mathbf{x} - \mathbf{a}\|) \end{cases} \quad (4.4)$$

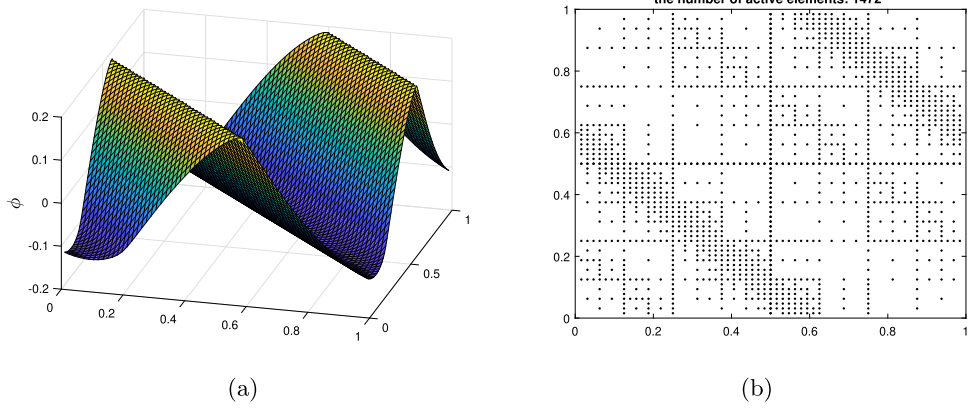
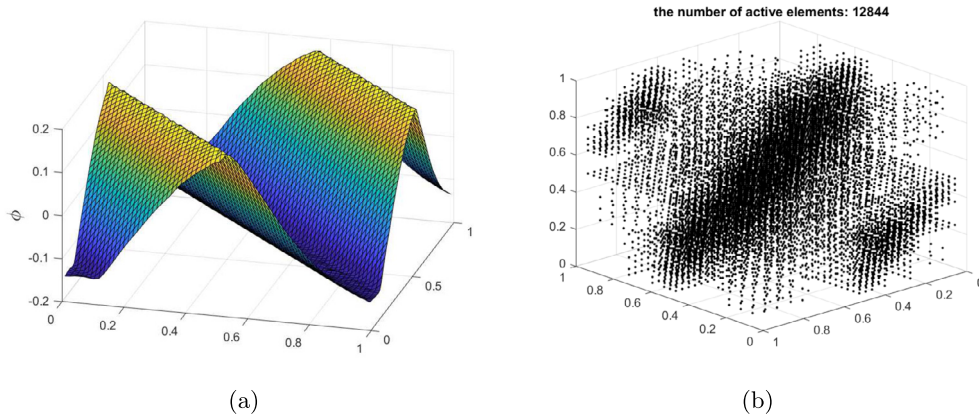
where $\mathbf{a} = (0.5, 0.5, \dots, 0.5)$ and

$$g(z) = \frac{1}{2r_0}(z^2 - r_0^2), \quad r_0 = \frac{1}{8}.$$

An outflow boundary condition is imposed. The viscosity solution is

Table 4.6Example 4.2, $d = 3$. Adaptive sparse grid. $T = 0.005$. $M = k$.

	ϵ	DoF	L_2 error	R_ϵ	R_{DoF}
$k = 1$	1.00E-03	2432	5.22E-03	–	–
	1.00E-04	15680	2.70E-03	0.29	0.35
	1.00E-05	46768	1.11E-03	0.39	0.82
	1.00E-06	151480	4.62E-04	0.38	0.74
	1.00E-07	391008	1.66E-04	0.45	1.08
$k = 2$	1.00E-03	2646	5.61E-03	–	–
	1.00E-04	8127	1.87E-03	0.48	0.98
	1.00E-05	21492	6.16E-04	0.48	1.15
	1.00E-06	55296	1.24E-04	0.70	1.70
	1.00E-07	154926	3.75E-05	0.52	1.16
$k = 3$	1.00E-03	1664	2.87E-03	–	–
	1.00E-04	6400	1.09E-03	0.42	0.72
	1.00E-05	17920	1.60E-04	0.83	1.87
	1.00E-06	44416	4.19E-05	0.58	1.47
	1.00E-07	186368	5.67E-06	0.87	1.40

**Fig. 4.5.** Example 4.2, $d = 2$. $T = 0.06$. $k = 2$, $M = k$. $N = 6$. $\epsilon = 10^{-5}$. (a) Numerical solutions on adaptive grids. (b) Active elements.**Fig. 4.6.** Example 4.2, $d = 3$. $k = 2$, $M = 3$. $T = 0.03$. $N = 6$. $\epsilon = 10^{-6}$. (a) 2D-cuts of the numerical solution at $x_3 = 0$. (b) Active elements.

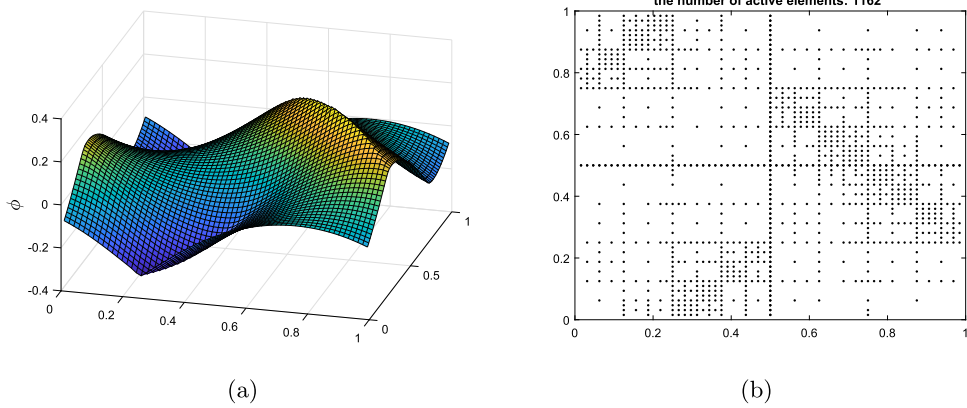
$$\phi(\mathbf{x}, t) = g(\max(\|\mathbf{x}\| - t, 0)),$$

which is clearly C^1 smooth.

One additional challenge of this problem is that the Hamiltonian is not smooth, making the DG formulation unstable if the numerical quadrature is not sufficiently accurate, as mentioned in previous section. To circumvent the difficulty, we propose to employ a regularized Hamiltonian as follows.

Table 4.7Example 4.3. Adaptive sparse grid. $T = 0.03$. $M = k$.

	ϵ	DoF	L_2 error	R_ϵ	R_{DoF}
$k = 1$	1.00E-03	180	9.78E-03	–	–
	1.00E-04	448	2.05E-03	0.68	1.71
	1.00E-05	952	1.26E-03	0.21	0.64
	1.00E-06	1296	2.24E-04	0.75	5.60
	1.00E-07	2952	2.74E-05	0.91	2.56
$k = 2$	1.00E-03	135	2.73E-03	–	–
	1.00E-04	306	3.95E-04	0.84	2.36
	1.00E-05	594	1.97E-04	0.30	1.05
	1.00E-06	1224	4.41E-05	0.65	2.07
	1.00E-07	2565	1.31E-05	0.53	1.64
$k = 3$	1.00E-03	112	5.38E-04	–	–
	1.00E-04	256	1.67E-04	0.51	1.42
	1.00E-05	560	4.29E-05	0.59	1.73
	1.00E-06	832	1.21E-05	0.55	3.20
	1.00E-07	1280	1.25E-06	0.99	5.27

**Fig. 4.7.** Example 4.3. $T = 0.2$. $N = 6$. $k = 2$, $M = 3$. $\epsilon = 10^{-5}$. (a) Numerical solution profile. (b) Active elements.**Table 4.8**Example 4.4, $d = 2, 3, 4$. Adaptive sparse grid. $T = 0.1$. $M = k + 1$.

	ϵ	$k = 1$				$k = 2$			
		DoF	L_2 error	R_ϵ	R_{DoF}	DoF	L_2 error	R_ϵ	R_{DoF}
$d = 2$	1.00E-03	236	2.25E-02			72	5.42E-03		
	1.00E-04	496	5.39E-03	0.62	1.92	108	3.59E-03	0.18	1.01
	1.00E-05	1056	2.93E-03	0.26	0.81	324	1.09E-03	0.52	1.09
	1.00E-06	1904	9.27E-04	0.50	1.95	900	4.41E-04	0.39	0.89
	1.00E-07	5496	2.43E-04	0.58	1.26	2880	1.23E-04	0.56	1.10
$d = 3$	1.00E-03	680	2.31E-02			108	6.46E-03		
	1.00E-04	1472	7.64E-03	0.48	1.43	351	3.18E-03	0.31	0.60
	1.00E-05	2968	4.12E-03	0.27	0.88	1026	1.53E-03	0.32	0.68
	1.00E-06	5080	1.72E-03	0.38	1.63	2970	5.76E-04	0.43	0.92
	1.00E-07	23272	5.00E-04	0.54	0.81	11610	2.26E-04	0.41	0.69
$d = 4$	1.00E-03	1872	2.30E-02			405	4.61E-03		
	1.00E-04	3792	2.52E-02	-0.04	-0.13	1053	2.90E-03	0.20	0.48
	1.00E-05	8944	1.25E-02	0.30	0.81	3159	1.16E-03	0.40	0.84
	1.00E-06	10624	2.59E-03	0.68	9.15	12312	5.71E-04	0.31	0.52
	1.00E-07	–	–	–	–	55080	2.03E-04	0.45	0.69

$$\tilde{H}(\nabla\phi) = \begin{cases} \|\nabla\phi\|, & \text{if } \|\nabla\phi\| \geq \delta \\ \frac{1}{2\delta}\|\nabla\phi\|^2 + \frac{1}{2}\delta, & \text{otherwise.} \end{cases} \quad (4.5)$$

It can be easily verified that \tilde{H} is C^1 . In the simulation, we choose $\delta = 2h$, where h is the mesh size, hence the regularization will not affect the accuracy of the original method. We employ the regularized Hamiltonian for all the tests, while we notice that it is only required for $k > 1$. In Table 4.8, we summarize the convergence study for the adaptive method for $d = 2, 3, 4$

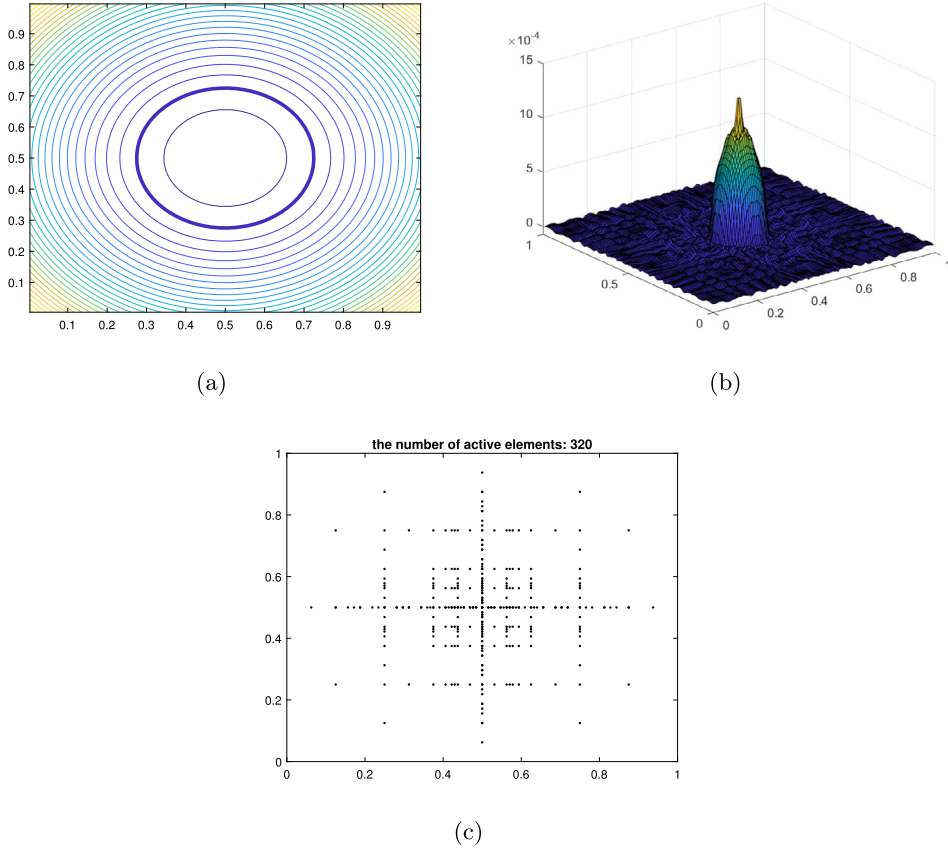


Fig. 4.8. Example 4.4, $d = 2$, $T = 0.1$, $k = 2$, $M = 3$, $N = 7$, $\epsilon = 10^{-7}$. (a) Contour plot of the numerical solution. (b) Numerical error distribution. (c) Active elements.

and $k = 1, 2$. It is observed that the convergence rates R_{DoF} and R_ϵ are similar for $k = 1$ and $k = 2$, which is unsurprising, since the viscosity solution is only C^1 . Meanwhile, the error magnitude by $k = 2$ is still much smaller than that by $k = 1$ with the same number of DoF, demonstrating the efficiency of method with high order accuracy. In Figs. 4.8-4.9, we report the contour plots of the numerical solutions with $N = 7$, $k = 2$, $M = 3$, $\epsilon = 10^{-7}$, $d = 2, 3$. For $d = 3$, we plot the 2D cuts at $x_1 = 0.61$ and at $x_2 = 0.51$. We observe that the rarefaction wave developed at the center of domain is correctly captured by the adaptive method. We also highlight the level set of $\phi = 0$ in the contour plots.

Example 4.5. In this example, we consider the following HJB equation [3]

$$\begin{cases} \phi_t + \max_{\mathbf{b} \in \mathcal{B}} \left(\sum_{m=1}^d b_m \cdot \nabla \phi \right) = 0, & \mathbf{x} \in [0, 1]^d, \\ \phi(\mathbf{x}, 0) = g(\|\mathbf{x} - \mathbf{a}\|), \end{cases} \quad (4.6)$$

where $\mathbf{a} = (0.5, 0.5, \dots, 0.5)$ and $\mathcal{B} = \{\mathbf{b} = (b_1, b_2, \dots, b_d), b_m = \pm 1\}$ is a set of 2^d vectors corresponding to 2^d possible controls. The function $g(z)$ is the same as in the Example 4.4. Note that this HJB equation is equivalent to the following HJ equation

$$\begin{cases} \phi_t + \sum_{m=1}^d |\phi_{x_m}| = 0, & \mathbf{x} \in [0, 1]^d \\ \phi(\mathbf{x}, 0) = g(\|\mathbf{x} - \mathbf{a}\|). \end{cases} \quad (4.7)$$

The exact solution can be hence derived from (4.7):

$$\phi(\mathbf{x}, t) = g(\|(\mathbf{x} - \mathbf{a})_t^*\|).$$

Here, for a vector \mathbf{c} , $\mathbf{c}_t^* := \min(\max(0, \mathbf{c} - t), \mathbf{c} + t)$ in the component-wise sense. We apply the adaptive algorithm to simulate (4.7). The outflow boundary conditions are imposed. Note that the Hamiltonian is nonsmooth as with the Eikonal equation,

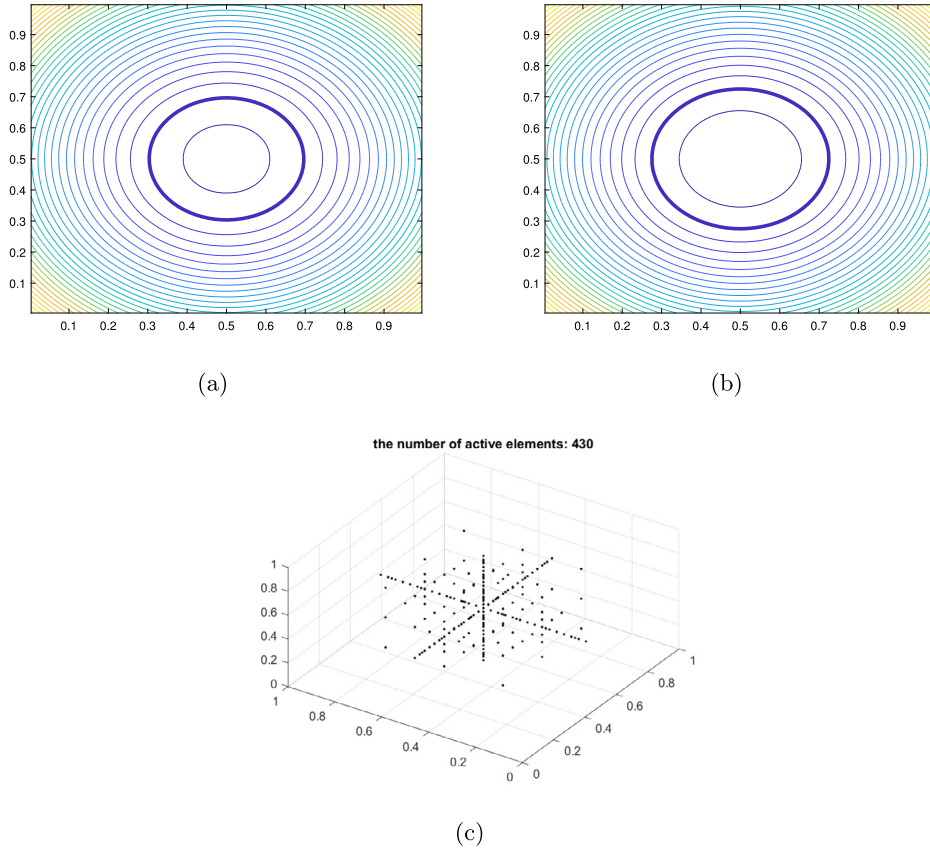


Fig. 4.9. Example 4.4, $d = 3$. $T = 0.1$. $k = 2$, $M = 3$. $N = 7$. $\epsilon = 10^{-7}$. (a) Contour plot of 2D cut at $x_1 = 0.61$. (b) Contour plot of 2D cut at $x_2 = 0.51$. (c) Active elements.

Table 4.9

Example 4.5, $d = 2, 3, 4$. Adaptive sparse grid. $T = 0.1$. $M = k + 2$.

ϵ	$k = 1$				$k = 2$			
	DoF	L_2 error	R_ϵ	R_{DoF}	DoF	L_2 error	R_ϵ	R_{DoF}
$d = 2$	1.00E-03	204	4.17E-03		63	8.62E-03		
	1.00E-04	444	1.62E-03	0.41	135	1.89E-03	0.66	1.99
	1.00E-05	860	7.01E-04	0.36	207	6.26E-04	0.48	2.59
	1.00E-06	876	6.43E-04	0.04	459	4.24E-04	0.17	0.49
	1.00E-07	924	6.58E-04	-0.01	855	3.97E-04	0.03	0.11
$d = 3$	1.00E-03	608	5.44E-03		270	1.12E-02		
	1.00E-04	1328	2.12E-03	0.41	594	2.98E-03	0.58	1.68
	1.00E-05	2576	9.15E-04	0.37	918	7.89E-04	0.58	3.05
	1.00E-06	2624	8.39E-04	0.04	2052	4.36E-04	0.26	0.74
	1.00E-07	2768	8.56E-04	-0.01	3510	3.93E-04	0.05	0.19
$d = 4$	1.00E-03	1616	6.65E-03		1053	1.35E-02		
	1.00E-04	3536	2.60E-03	0.41	1701	6.17E-03	0.34	1.63
	1.00E-05	6864	1.12E-03	0.37	2997	1.24E-03	0.70	2.84
	1.00E-06	6992	1.02E-03	0.04	6237	4.90E-04	0.40	1.26
	1.00E-07	7376	1.05E-03	-0.01	12069	4.20E-04	0.07	0.23

and hence we regularize the absolute function using the technique (4.5) to ensure stability. In Fig. 4.10, we plot the solution with configuration $k = 2$, $M = 4$, $N = 7$, $\epsilon = 10^{-7}$. Note that the viscosity solution is C^1 , a rarefaction wave opens up at the center of the domain, which is well captured by the method. In Table 4.9, we summarize the convergence study for $d = 2, 3, 4$ and $k = 1, 2$. Note that when $\epsilon = 10^{-7}$, the error does not decay anymore, since it has saturated already with the maximum level $N = 7$.

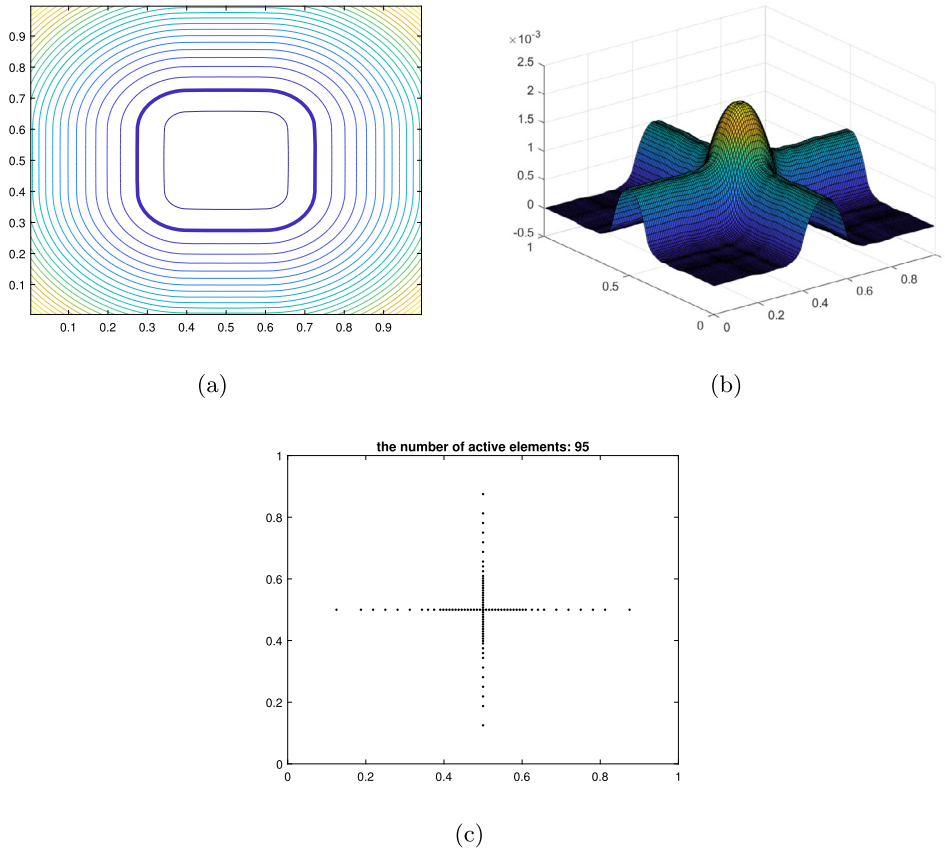


Fig. 4.10. Example 4.5. $T = 0.1$, $k = 2$, $M = 4$, $N = 7$, $\epsilon = 10^{-7}$. (a) Contour plot of the numerical solution. (b) Numerical error distribution. (c) Active elements.

Example 4.6. In the last example, we consider the 2D problem related to controlling optimal cost determination [39]

$$\begin{cases} \phi_t - \sin(2\pi x_2)\phi_{x_1} - (\sin(2\pi x_1) + \text{sign}(\phi_{x_2}))\phi_{x_2} - \frac{1}{2}\sin^2(2\pi x_2) - \cos(2\pi x_1) - 1 = 0, & \mathbf{x} \in [0, 1]^2 \\ \phi(\mathbf{x}, 0) = 0 \end{cases} \quad (4.8)$$

Note that the Hamiltonian is not smooth. In Fig. 4.11, we plot the solution profile, the optimal $\text{sign}(\phi_{x_2})$ together with the active elements at final time $T = 0.15$. Again, we regularize the Hamiltonian as with previous examples. The adaptive method is able to capture the viscosity solution efficiently, and the numerical results agree with other methods in the literature, e.g. [24,34,6,22,30].

5. Conclusion

In this work, we proposed an adaptive sparse grid LDG method for solving HJ equations in high dimensions. By incorporating the orthonormal Alpert's multiwavelets as the DG finite element bases, and the interpolatory multiwavelets as efficient multiresolution numerical quadratures, we achieve efficient multiresolution schemes which is suitable for high dimensions. Benchmark numerical tests up to 4D are provided to validate the performance of the method. The code generating the results in this paper can be found at the GitHub link: <https://github.com/JuntaoHuang/adaptive-multiresolution-DG>, and it has the capability of computing higher dimensional problems.

CRediT authorship contribution statement

Wei Guo: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Juntao Huang:** Conceptualization, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Zhanjing Tao:** Conceptualization, Methodology, Writing – review & editing. **Yingda Cheng:** Conceptualization, Methodology, Writing – review & editing.

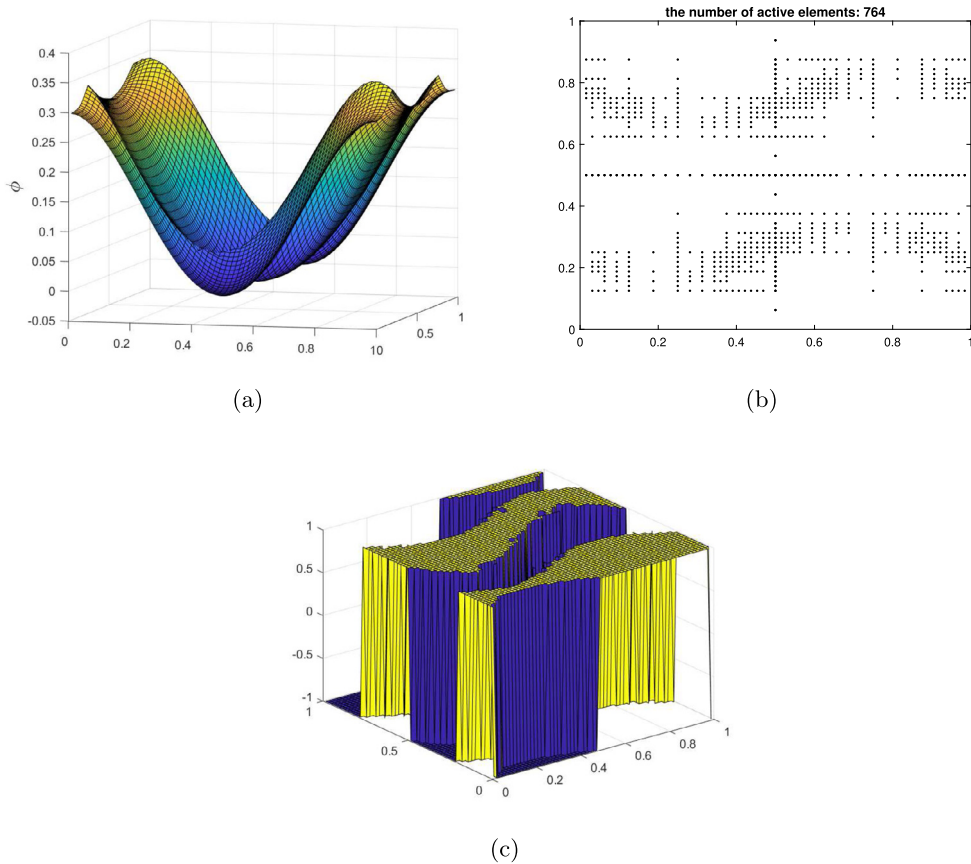


Fig. 4.11. Example 4.6, $d = 2$, $T = 0.15$, $k = 2$, $M = 4$, $N = 6$, $\epsilon = 10^{-5}$. (a) Numerical solution profile. (b) Active elements. (c) Controls $\text{sign}(\phi_y)$.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Qi Tang and Kai Huang for the assistance and discussion in code implementation. Yingda Cheng would like to thank the support from IPAM to attend the workshop on “High-dimensional Hamilton-Jacobi PDEs”.

References

- [1] R. Abgrall, Numerical discretization of the first-order Hamilton-Jacobi equation on triangular meshes, *Commun. Pure Appl. Math.* 49 (12) (1996) 1339–1373.
- [2] B. Alpert, A class of bases in L^2 for the sparse representation of integral operators, *SIAM J. Math. Anal.* 24 (1) (1993) 246–262.
- [3] O. Bokanowski, J. Garcke, M. Griebel, I. Klompaker, An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations, *J. Sci. Comput.* 55 (3) (2013) 575–605.
- [4] H.-J. Bungartz, M. Griebel, Sparse grids, *Acta Numer.* 13 (2004) 147–269.
- [5] Y. Cheng, C.-W. Shu, A discontinuous Galerkin finite element method for directly solving the Hamilton-Jacobi equations, *J. Comput. Phys.* 223 (1) (2007) 398–415.
- [6] Y. Cheng, Z. Wang, A new discontinuous Galerkin finite element method for directly solving the Hamilton-Jacobi equations, *J. Comput. Phys.* 268 (2014) 134–153.
- [7] Y. Chow, J. Darbon, S. Osher, W. Yin, Algorithm for overcoming the curse of dimensionality for time-dependent non-convex Hamilton-Jacobi equations arising from optimal control and differential games problems, *J. Sci. Comput.* 73 (2–3) (2017) 617–643.
- [8] Y.T. Chow, J. Darbon, S. Osher, W. Yin, Algorithm for overcoming the curse of dimensionality for state-dependent Hamilton-Jacobi equations, *J. Comput. Phys.* 387 (2019) 376–409.
- [9] B. Cockburn, S. Hou, C.-W. Shu, The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multi-dimensional case, *Math. Comput.* 54 (190) (1990) 545–581.
- [10] B. Cockburn, C.-W. Shu, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, *SIAM J. Numer. Anal.* 35 (6) (1998) 2440–2463.

- [11] M.G. Crandall, L.C. Evans, P.-L. Lions, Some properties of viscosity solutions of Hamilton-Jacobi equations, *Trans. Am. Math. Soc.* 282 (2) (1984) 487–502.
- [12] M.G. Crandall, P.-L. Lions, Viscosity solutions of Hamilton-Jacobi equations, *Trans. Am. Math. Soc.* 277 (1) (1983) 1–42.
- [13] M.G. Crandall, P.L. Lions, Two approximations of solutions of Hamilton-Jacobi equations, *Math. Comput.* 43 (167) (1984) 1–19.
- [14] J. Darbon, G.P. Langlois, T. Meng, Overcoming the curse of dimensionality for some Hamilton-Jacobi partial differential equations via neural network architectures, *arXiv preprint*, arXiv:1910.09045, 2019.
- [15] J. Darbon, T. Meng, On some neural network architectures that can represent viscosity solutions of certain high dimensional Hamilton-Jacobi partial differential equations, *arXiv preprint*, arXiv:2002.09750, 2020.
- [16] J. Darbon, S. Osher, Algorithms for overcoming the curse of dimensionality for certain Hamilton-Jacobi equations arising in control theory and elsewhere, *Res. Math. Sci.* 3 (1) (2016) 19.
- [17] S. Dolgov, D. Kalise, K. Kunisch, Tensor decompositions for high-dimensional Hamilton-Jacobi-Bellman equations, *arXiv preprint*, arXiv:1908.01533, 2019.
- [18] L. Evans, *Partial Differential Equations*, second edition, American Mathematical Society, 2010.
- [19] J. Garcke, A. Kröner, Suboptimal feedback control of PDEs by solving HJB equations on adaptive sparse grids, *J. Sci. Comput.* 70 (1) (2017) 1–28.
- [20] W. Guo, Y. Cheng, A sparse grid discontinuous Galerkin method for high-dimensional transport equations and its application to kinetic simulations, *SIAM J. Sci. Comput.* 38 (6) (2016) A3381–A3409.
- [21] W. Guo, Y. Cheng, An adaptive multiresolution discontinuous Galerkin method for time-dependent transport equations in multidimensions, *SIAM J. Sci. Comput.* 39 (6) (2017) A2962–A2992.
- [22] W. Guo, F. Li, J. Qiu, Local-structure-preserving discontinuous Galerkin methods with Lax-Wendroff type time discretizations for Hamilton-Jacobi equations, *J. Sci. Comput.* 47 (2) (2011) 239–257.
- [23] J. Han, A. Jentzen, E. Weinan, Solving high-dimensional partial differential equations using deep learning, *Proc. Natl. Acad. Sci.* 115 (34) (2018) 8505–8510.
- [24] C. Hu, C.-W. Shu, A discontinuous Galerkin finite element method for Hamilton-Jacobi equations, *SIAM J. Sci. Comput.* 21 (2) (1999) 666–690.
- [25] J. Huang, Y. Cheng, An adaptive multiresolution discontinuous Galerkin method with artificial viscosity for scalar hyperbolic conservation laws in multidimensions, *arXiv preprint*, arXiv:1906.00829, 2019.
- [26] J. Huang, Y. Liu, W. Guo, Z. Tao, Y. Cheng, An adaptive multiresolution interior penalty discontinuous Galerkin method for wave equations in second order form, *arXiv preprint*, arXiv:2004.08525, 2020.
- [27] J. Huang, C.-W. Shu, Error estimates to smooth solutions of semi-discrete discontinuous Galerkin methods with quadrature rules for scalar conservation laws, *Numer. Methods Partial Differ. Equ.* 33 (2) (2017) 467–488.
- [28] G.-S. Jiang, D. Peng, Weighted ENO schemes for Hamilton-Jacobi equations, *SIAM J. Sci. Comput.* 21 (6) (2000) 2126–2143.
- [29] W. Kang, L.C. Wilcox, Mitigating the curse of dimensionality: sparse grid characteristics method for optimal feedback control and HJB equations, *Comput. Optim. Appl.* 68 (2) (2017) 289–315.
- [30] G. Ke, W. Guo, An alternative formulation of discontinuous Galerkin schemes for solving Hamilton-Jacobi equations, *J. Sci. Comput.* 78 (2) (2019) 1023–1044.
- [31] K. Kunisch, S. Volkwein, L. Xie, HJB-POD-based feedback design for the optimal control of evolution problems, *SIAM J. Appl. Dyn. Syst.* 3 (4) (2004) 701–722.
- [32] F. Lafon, S. Osher, High order two dimensional nonoscillatory methods for solving Hamilton-Jacobi scalar equations, *J. Comput. Phys.* 123 (2) (1996) 235–253.
- [33] O. Lepsky, C. Hu, C.-W. Shu, Analysis of the discontinuous Galerkin method for Hamilton-Jacobi equations, *Appl. Numer. Math.* 33 (1–4) (2000) 423–434.
- [34] F. Li, S. Yakovlev, A central discontinuous Galerkin method for Hamilton-Jacobi equations, *J. Sci. Comput.* 45 (1–3) (2010) 404–428.
- [35] P.-L. Lions, *Generalized Solutions of Hamilton-Jacobi Equations*, vol. 69, Pitman, London, 1982.
- [36] S. Mallat, *A Wavelet Tour of Signal Processing*, Elsevier, 1999.
- [37] T. Nakamura-Zimmerer, Q. Gong, W. Kang, Adaptive deep learning for high dimensional Hamilton-Jacobi-Bellman equations, *arXiv preprint*, arXiv:1907.05317, 2019.
- [38] S. Osher, J. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, *J. Comput. Phys.* 79 (1) (1988) 12–49.
- [39] S. Osher, C.-W. Shu, High-order essentially nonoscillatory schemes for Hamilton-Jacobi equations, *SIAM J. Numer. Anal.* 28 (4) (1991) 907–922.
- [40] J. Qiu, C.-W. Shu, Hermite WENO schemes for Hamilton-Jacobi equations, *J. Comput. Phys.* 204 (1) (2005) 82–99.
- [41] J. Shen, H. Yu, Efficient spectral sparse grid methods and applications to high-dimensional elliptic problems, *SIAM J. Sci. Comput.* 32 (6) (2010) 3228–3250.
- [42] C.-W. Shu, High order numerical methods for time dependent Hamilton-Jacobi equations, in: *Mathematics and Computation in Imaging Science and Information Processing*, in: *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.*, vol. 11, 2007, pp. 47–91.
- [43] C.-W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock-capturing schemes, *J. Comput. Phys.* 77 (2) (1988) 439–471.
- [44] Z. Tao, Y. Jiang, Y. Cheng, An adaptive high-order piecewise polynomial based sparse grid collocation method with applications, *arXiv preprint*, arXiv:1912.03982, 2019.
- [45] Z. Wang, Q. Tang, W. Guo, Y. Cheng, Sparse grid discontinuous Galerkin methods for high-dimensional elliptic equations, *J. Comput. Phys.* 314 (2016) 244–263.
- [46] J. Yan, S. Osher, A local discontinuous Galerkin method for directly solving Hamilton-Jacobi equations, *J. Comput. Phys.* 230 (1) (2011) 232–244.
- [47] Y.-T. Zhang, C.-W. Shu, High-order WENO schemes for Hamilton-Jacobi equations on triangular meshes, *SIAM J. Sci. Comput.* 24 (3) (2003) 1005–1030.