Kinetic Models and High Dimensional Scientific Computing

Yingda Cheng

Virginia Tech

Applied Numerical Analysis Seminar, Sept. 2023

Outline

1. Part I: Kinetic models

2. Part II: High dimensional scientific computing
   - Sparse grid DG method
   - Adaptivity and nonlinearity
   - Quick summary

Outline

Part I: Kinetic Models

Goal: to introduce a class of models that are important in physics.

Disclamer: *this is not my work, but from classical literature. The materials are from physics and math, and the content can be dense!*

In the end, I hope you find the models and their structures interesting.

Multiscale Models

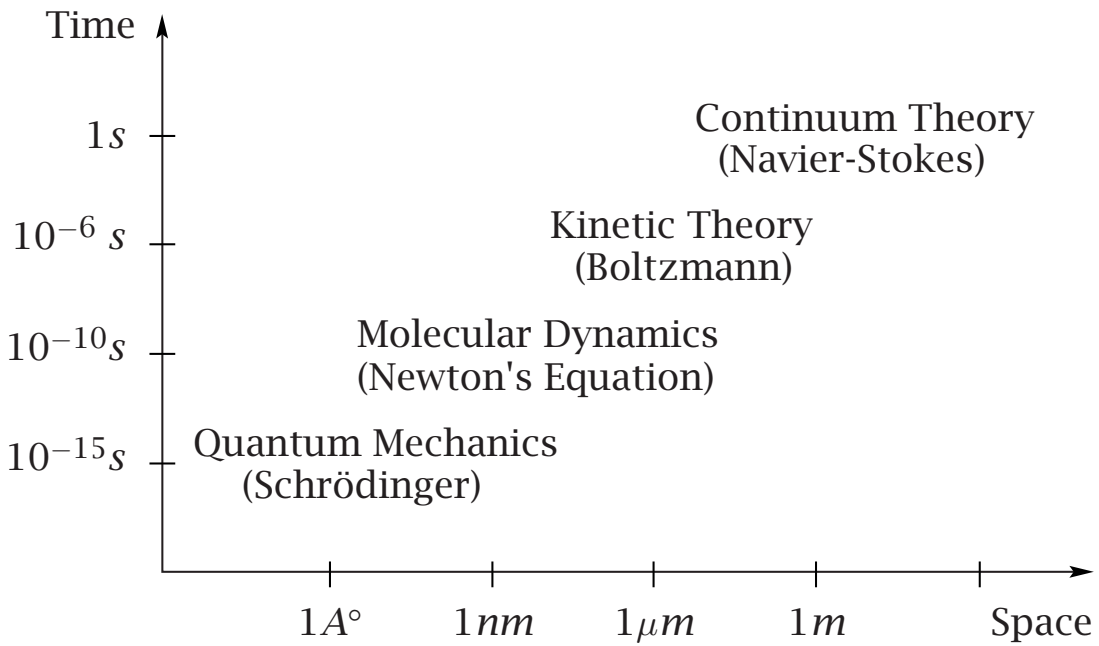


Figure: Figure from E & Engquist, AMS Notice, 2003

Kinetic model provide a *mesoscopic description* of interacting particle system, and is the key for multiscale modeling connecting MD & Fluid equations.

Boltzmann Transport Equations (BTE)

Boltzmann equation

Article Talk

From Wikipedia, the free encyclopedia

For other uses, see Boltzmann's entropy formula, Stefan–Boltzmann law, and Maxwell–Boltzmann distribution. "BTE" redirects here. For other uses, see BTE (disambiguation).

The **Boltzmann equation** or **Boltzmann transport equation** (**BTE**) describes the statistical behaviour of a thermodynamic system not in a state of equilibrium; it was devised by Ludwig Boltzmann in 1872.[2] The classic example of such a system is a fluid with temperature gradients in space causing heat to flow from hotter regions to colder ones, by the random but biased transport of the particles making up that fluid. In the modern literature the term Boltzmann equation is often used in a more general sense, referring to any kinetic equation that describes the change of a macroscopic quantity in a thermodynamic system, such as energy, charge or particle number.

Boltzmann Transport Equations (BTE)

Boltzmann equation

Article Talk

From Wikipedia, the free encyclopedia

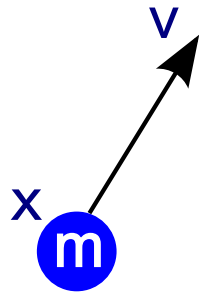
For other uses, see Boltzmann's entropy formula, Stefan–Boltzmann law, and Maxwell–Boltzmann distribution.
"BTE" redirects here. For other uses, see BTE (disambiguation).

The **Boltzmann equation** or **Boltzmann transport equation** (**BTE**) describes the statistical behaviour of a thermodynamic system not in a state of equilibrium; it was devised by Ludwig Boltzmann in 1872.[2] The classic example of such a system is a fluid with temperature gradients in space causing heat to flow from hotter regions to colder ones, by the random but biased transport of the particles making up that fluid. In the modern literature the term Boltzmann equation is often used in a more general sense, referring to any kinetic equation that describes the change of a macroscopic quantity in a thermodynamic system, such as energy, charge or particle number.

Next, I will describe 3 classical examples of kinetic models.

Example I: Classical Boltzmann from Rarefied Gas

Equations of motion



$$\begin{cases} \frac{d\mathbf{x}_i}{dt} = \mathbf{v}_i \\ \frac{d\mathbf{v}_i}{dt} = \frac{\mathbf{F}_i}{m} \end{cases}$$

where $\mathbf{F}_i$ accounts for external force and particle interactions (e.g. binary collision terms).

For a N-particle system, we will have the state vector $(\mathbf{x}_1, \ldots \mathbf{x}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N)$. This has $6N$ unknowns, and in realistic case $N \approx 10^{20}$.

Statistical description to the rescue

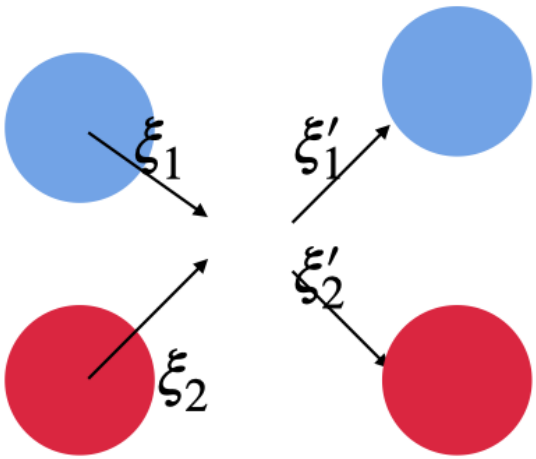
The Boltzmann equation considers the **Probability Density Function (pdf)** $f(t, \mathbf{x}, \mathbf{v})$, where $f(t, \mathbf{x}, \mathbf{v}) d\mathbf{x} d\mathbf{v}$ describes the probability of finding one particle in an infinitesimal volume $d\mathbf{x} d\mathbf{v}$ centered at the point $(\mathbf{x}, \mathbf{v})$ of the phase space. BTE reads

$$\frac{Df}{Dt} = \underbrace{\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{x}} f}_{transport} = \underbrace{Q(f, f)}_{collision}$$

Illustration: Microscopic process[1]

Note: we switched $\mathbf{v}$ to ξ.

Microscopic process
Binary collisions



Conservation

$$\xi_1 + \xi_2 = \xi_1' + \xi_2'$$

$$|\xi_1|^2 + |\xi_2|^2 = |\xi_1'|^2 + |\xi_2'|^2$$

[1] C. Cercignani. Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations. Cambridge University Press, Cambridge, 2000.

Derivation[2]

We consider the marginal distribution $P^{(1)}(x_1, \xi_1, t)$, which is probability of finding particle 1 at (x_1, ξ_1, t).

$$\frac{\partial P^{(1)}}{\partial t} + \boldsymbol{\xi}_1 \cdot \frac{\partial P^{(1)}}{\partial \mathbf{x}_1} = G - L.$$

$$G - L = (N-1)\sigma^2 \int_{R^3} \int_{\mathcal{B}} P^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \sigma \mathbf{n}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, t)(\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1) \cdot \mathbf{n} d\boldsymbol{\xi}_2 d\mathbf{n}, \tag{1.2.8}$$

[2] C. Cercignani. Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations. Cambridge University Press, Cambridge, 2000.

Derivation[2]

We consider the marginal distribution $P^{(1)}(x_1, \xi_1, t)$, which is probability of finding particle 1 at (x_1, ξ_1, t).

$$\frac{\partial P^{(1)}}{\partial t} + \boldsymbol{\xi}_1 \cdot \frac{\partial P^{(1)}}{\partial \mathbf{x}_1} = G - L.$$

$$G - L = (N-1)\sigma^2 \int_{R^3} \int_{\mathcal{B}} P^{(2)}(\mathbf{x}_1, \mathbf{x}_1 + \sigma \mathbf{n}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, t)(\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1) \cdot \mathbf{n} d\boldsymbol{\xi}_2 d\mathbf{n}, \tag{1.2.8}$$

By **molecular chaos assumption**
$P^{(2)}(x_1, \xi_1, x_2, \xi_2, t) = P^{(1)}(x_1, \xi_1, t) P^{(1)}(x_2, \xi_2, t)$, then

$$\frac{\partial P^{(1)}}{\partial t} + \boldsymbol{\xi}_1 \cdot \frac{\partial P^{(1)}}{\partial \mathbf{x}_1} = N\sigma^2 \int_{R^3} \int_{\mathcal{B}^-} \left[P^{(1)}(\mathbf{x}_1, \boldsymbol{\xi}_1', t) P^{(1)}(\mathbf{x}_1, \boldsymbol{\xi}_2', t) \right.$$
$$\left. - P^{(1)}(\mathbf{x}_1, \boldsymbol{\xi}_1, t) P^{(1)}(\mathbf{x}_1, \boldsymbol{\xi}_2, t) \right] |(\boldsymbol{\xi}_2 - \boldsymbol{\xi}_1) \cdot \mathbf{n}| d\boldsymbol{\xi}_2 d\mathbf{n}. \tag{1.2.16}$$

[2] C. Cercignani. Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations. Cambridge University Press, Cambridge, 2000.

Properties

Now $P^{(1)}$ is f, and we have

$$\frac{\partial f}{\partial t} + \boldsymbol{\xi} \cdot \frac{\partial f}{\partial \mathbf{x}} = \int_{R^3} \int_{\mathcal{B}^-} (f' f'_* - f f_*) \mathcal{B}(\theta, V) d\boldsymbol{\xi}_* d\theta d\epsilon,$$

The right hand side $Q(f, f)$ models the binary particle collision, and satisfies the following properties.

- $\int Q(f, f) d\xi = 0$ mass conservation.
- $\int Q(f, f) \xi dv = 0$ momentum conservation.
- $\int Q(f, f) |\xi|^2 dv = 0$ energy conservation.
- $\int Q(f, f) log(f) dv \leq 0$ Boltzmann H-theorem. Further $Q(f, f) = 0$ iff. f is a Maxwellian distribution. (statistical equilibrium)

Bridging the scales

- We have derived the original version of the BTE from microscopic physical laws. (Model reduction from **micro to meso**).
- Now, we introduce scaling $Kn = l/d = \epsilon$.

$$f_t + \xi \cdot \nabla_x f = \frac{1}{\epsilon} Q(f, f).$$

  As $\epsilon \to 0$, this is dense gas. We can derive the Euler/NS equations through Chapman-Enskog expansions of the BTE. (Model reduction from **meso to macro**).
- $\epsilon \to \infty$, particle free flow. (This is **meso** scale, hard to reduce.)

Example II: Radiation transport

- So far, we have derived the original version of the BTE. It looks too complicated. As mathematicians, we like to simplify.

Example II: Radiation transport

- So far, we have derived the original version of the BTE. It looks too complicated. As mathematicians, we like to simplify.
- Perhaps, the most famous simplified version is the BGK model (a nonlinear relaxation model).

Example II: Radiation transport

- So far, we have derived the original version of the BTE. It looks too complicated. As mathematicians, we like to simplify.
- Perhaps, the most famous simplified version is the BGK model (a nonlinear relaxation model).
- But here let's consider an even simpler model: a **linear** relaxation model.

$$v \in [-1,1], \quad \epsilon f_t + v f_x = \frac{\rho - f}{\epsilon}, \quad \text{where} \quad \rho = \int_{-1}^{1} f dv$$

- With macro-micro decomposition $f = f_0 + \epsilon g$, when $\epsilon \to 0$, $f_0 = \rho, g = -v\rho_x$ with $\rho_t = -\int_{-1}^{1} v g_x dv = \frac{1}{3}\rho_{xx}$. Heat equation.

Example II: radiative transfer

That looks too simple. Let's consider a similar but 'real' problem from nuclear engineering: the neutron transport [3]

> **Neutron transport** (also known as **neutronics**) is the study of the motions and interactions of neutrons with materials. Nuclear scientists and engineers often need to know where neutrons are in an apparatus, in what direction they are going, and how quickly they are moving. It is commonly used to determine the behavior of nuclear reactor cores and experimental or industrial neutron beams. Neutron transport is a type of radiative transport.

Linear equation: steady state version

$$\boldsymbol{\Omega} \cdot \nabla_{\mathbf{x}} \varphi + \Sigma_t(\mathbf{x})\varphi = \Sigma_s(\mathbf{x})(\mathcal{S}\varphi)(\mathbf{x}) + Q(\mathbf{x}), \quad \mathbf{x} \in \Omega_{\mathbf{x}},\ \boldsymbol{\Omega} \in \mathbb{S}^{d-1}, \tag{1}$$

where the nonnegative $\Sigma_s(\mathbf{x})$, $\Sigma_a(\mathbf{x})$ and $\Sigma_t = \Sigma_s + \Sigma_a$, respectively, are the scattering, absorption and total cross sections. $Q(\mathbf{x})$ is an external source. $\mathcal{S}\varphi = \langle\varphi\rangle := \frac{1}{|\mathbb{S}^{d-1}|}\int_{\mathbb{S}^{d-1}} \varphi(\cdot, \boldsymbol{\Omega}) d\boldsymbol{\Omega}$.

[3] Lewis, Elmer Eugene, and Warren F. Miller. Computational methods of neutron transport. (1984).

Example II

- If Σ_s is large, this is optically thick region, more towards the diffusion limit.

Example II

- If Σ_s is large, this is optically thick region, more towards the diffusion limit.
- We can have device configuration with largely varying and discontinuous scattering coefficient.

Example II

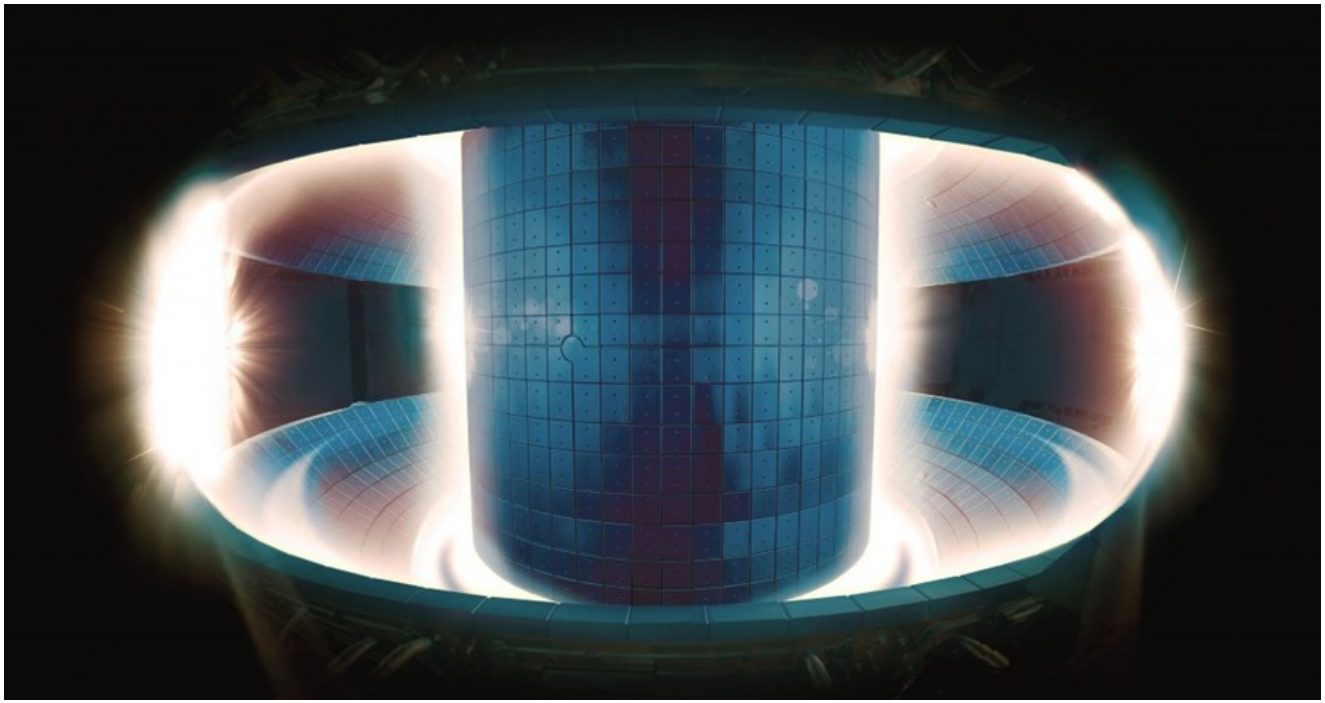
- If Σ_s is large, this is optically thick region, more towards the diffusion limit.
- We can have device configuration with largely varying and discontinuous scattering coefficient.
- Alternative formulation: time dependent RTE, and also eigenvalue problem for nuclear reactor criticality.

Example II

- If Σ_s is large, this is optically thick region, more towards the diffusion limit.
- We can have device configuration with largely varying and discontinuous scattering coefficient.
- Alternative formulation: time dependent RTE, and also eigenvalue problem for nuclear reactor criticality.
- This model is also used in astrophysics and in optical tomography. Understanding thermal radiative transfer is key to the inertial confinement fusion (ICF).

Example III: Vlasov dynamics in collisionless plasma

Now we switch gear to another **nonlinear** kinetic system. The Vlasov equation is the key to understanding kinetic effects in plasmas [4].



[4] Bittencourt, J. A. (2004). Fundamentals of plasma physics. Springer Science & Business Media

The one-species Vlasov-Maxwell (VM) system

The single-species VM system under the scaling of the characteristic time by the inverse of the plasma frequency ω_p^{-1} and length scaled by the Debye length λ_D, and characteristic electric and magnetic field as $\bar{E} = \bar{B} = -mc\omega_p/e$ is

$$\begin{aligned}
&\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + (\mathbf{E} + \mathbf{v}\times\mathbf{B})\cdot\nabla_{\mathbf{v}} f = 0\ , \\
&\frac{\partial \mathbf{E}}{\partial t} = \nabla_{\mathbf{x}}\times\mathbf{B} - \mathbf{J}, \qquad \frac{\partial \mathbf{B}}{\partial t} = -\nabla_{\mathbf{x}}\times\mathbf{E}\ , \\
&\nabla_{\mathbf{x}}\cdot\mathbf{E} = \rho - \rho_i, \qquad \nabla_{\mathbf{x}}\cdot\mathbf{B} = 0\ ,
\end{aligned}$$

where the density and current density are defined as

$$\rho(\mathbf{x},t) = \int_{\mathbb{R}^n} f(\mathbf{x},\mathbf{v},t)d\mathbf{v}, \qquad \mathbf{J}(\mathbf{x},t) = \int_{\mathbb{R}^n} f(\mathbf{x},\mathbf{v},t)\mathbf{v}d\mathbf{v}.$$

and ρ_i is the ion density.

The Vlasov-Ampère (VA) and Vlasov-Poisson (VP) system

In the zero-magnetic limit, the VM system becomes

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{E} \cdot \nabla_{\mathbf{v}} f = 0 \,, \tag{2}$$
$$\frac{\partial \mathbf{E}}{\partial t} = -\mathbf{J}, \qquad \nabla_{\mathbf{x}} \cdot \mathbf{E} = \rho - \rho_i,$$

This leads to either the Vlasov-Ampère (VA) system

$$\partial_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f + \mathbf{E} \cdot \nabla_{\mathbf{v}} f = 0 \,, \tag{3}$$
$$\frac{\partial \mathbf{E}}{\partial t} = -\mathbf{J},$$

The Vlasov-Ampère (VA) and Vlasov-Poisson (VP) system

In the zero-magnetic limit, the VM system becomes

$$\begin{aligned} &\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \mathbf{E}\cdot\nabla_{\mathbf{v}} f = 0\ , \\ &\frac{\partial \mathbf{E}}{\partial t} = -\mathbf{J}, \qquad \nabla_{\mathbf{x}}\cdot\mathbf{E} = \rho - \rho_i, \end{aligned} \tag{2}$$

This leads to either the Vlasov-Ampère (VA) system

$$\begin{aligned} &\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \mathbf{E}\cdot\nabla_{\mathbf{v}} f = 0\ , \\ &\frac{\partial \mathbf{E}}{\partial t} = -\mathbf{J}, \end{aligned} \tag{3}$$

or the Vlasov-Poisson (VP) system

$$\begin{aligned} &\partial_t f + \mathbf{v}\cdot\nabla_{\mathbf{x}} f + \mathbf{E}\cdot\nabla_{\mathbf{v}} f = 0\ , \\ &\nabla_{\mathbf{x}}\cdot\mathbf{E} = \rho - \rho_i, \end{aligned} \tag{4}$$

They are equivalent when the charge continuity equation

$$\rho_t + \nabla_{\mathbf{x}}\cdot\mathbf{J} = 0$$

is satisfied. (No external field)

A little bit about the models

- Mass and energy $\int fv^2 dxdv + \int (E^2 + B^2)dx$ is preserved. The system is Hamiltonian.

A little bit about the models

- Mass and energy $\int fv^2 dxdv + \int (E^2 + B^2)dx$ is preserved. The system is Hamiltonian.
- To understand VP system, let's simplify and consider particle free streaming, $f_t + vf_x = 0$.
  The exact solution is $f(t, x, v) = f_0(x - vt, v)$. Therefore, $|\frac{\partial f}{\partial v}| \approx t$ increase with time.
- Without collision, this gives rise to 'filamentation'. Thin filaments will occur in the phase space over time. Eventually, the solvers all run out of resolution.

A little bit about the models

- Mass and energy $\int fv^2 dxdv + \int (E^2 + B^2)dx$ is preserved. The system is Hamiltonian.
- To understand VP system, let's simplify and consider particle free streaming, $f_t + vf_x = 0$.
  The exact solution is $f(t, x, v) = f_0(x - vt, v)$. Therefore, $|\frac{\partial f}{\partial v}| \approx t$ increase with time.
- Without collision, this gives rise to 'filamentation'. Thin filaments will occur in the phase space over time. Eventually, the solvers all run out of resolution.
- The relevant collisional term is from electron scattering, e.g. Coulomb collisions. This gives rise to the Landau-Fokker-Planck equation. This is the fundamental model in magnetic confinement fusion (MCF).
- Nonlinearity gives rise to many interesting physics.

Summary

- BTE models the PDF (this is a **high dimensional** function)
- There are two parts of BTE: transport and collision.
- In some regime ($\epsilon \to 0$, collision dominated), we recover macroscopic equations by rigorous/heuristic argument. **Analytic model reduction** In other regime (big ϵ), we can have transport dominated case. The whole thing is **Multiscale**
- There are rich structures for the solution (conservation, asymptotic limit, Hamiltonian...).
- The modeling and simulations are particularly relevant to national lab (nuclear reactor, fusion energy). Applications not mentioned: semiconductor device, astrophysics, social dynamics, etc.

Summary on numerics

- Two classes of methods: deterministic and probabilistic.
- Probabilistic methods include DSMC, particle in cell. Solutions are noisy.
- Deterministic methods (i.e. PDE solvers). For gas Boltzmann, fast spectral method. RTE: DG in $\mathbf{x}$ and discrete coordinate in angles. Vlasov: semi-Lagrangian, DG etc.
- **Computational challenges:** need to observe physical laws (conservation), intrinsic multiscale behavior (the model can span several regimes, e.g. from transport dominated to diffusion dominated regimes).
- **Main computational challenge:** high dimensions.

Outline

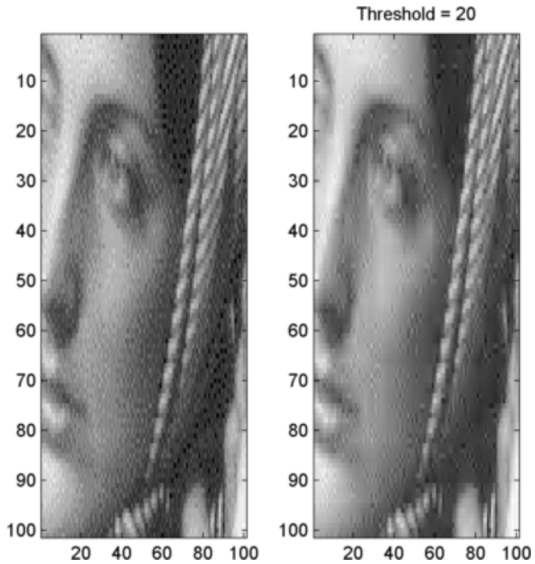
Overview

- Kinetic models are in high dimensions.
- Other high dimensional models: Hamilton-Jacobi-Bellman (HJB) from control, Schröndinger equation from quantum dynamics.
- Parametric PDEs can be of high dimension.
- Even for lower dimensional PDEs, we can 'unfold' the complexity into higher dimensions.

High D problem is hard. DOF scales like $O(N^d)$, fixed order error is $O(N^{-k})$, therefore error behaves like $O(DOF^{-k/d})$. **No storage, No accuracy!**

Wavelet compression

- We are interested in solving high dimensional PDEs, and developing numerical solvers that has stability and accuracy (if possible).
- The main tools we use are multiresolution analysis (MRA).
- Wavelet compression is widely used in signal and image processing.



Sparse grid: a tool to break the curse of dimensionality

- Sparse grid method is introduced by Smolyak (63) for high dimensional quadrature, and widely used for uncertainty quantification Xiu, Hesthaven (05...).
- Sparse grid PDE solver: Zenger (91), Griebel (91,98,05...). Most work focus on continuous FEM, and spectral methods.

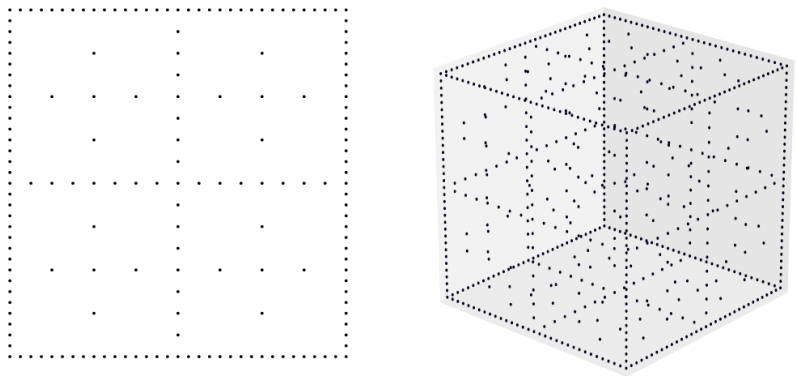
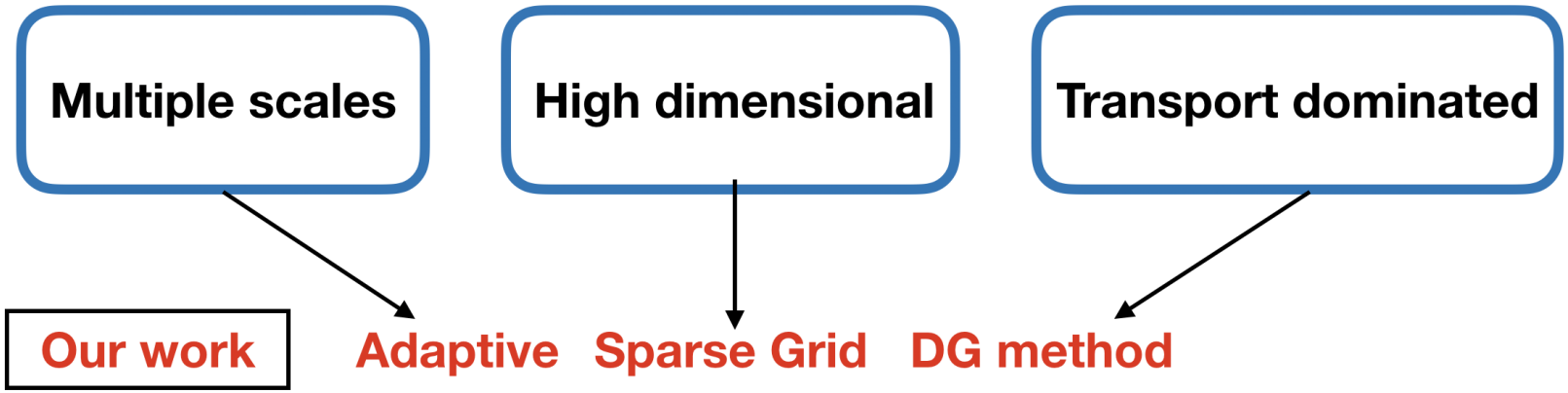


Fig. 5: Two-dimensional sparse grid (left) and three-dimensional sparse grid (right) of level $n = 5$.

Figure: From Garcke, SG in a nutshell

Our approach



Our work & outline for this section

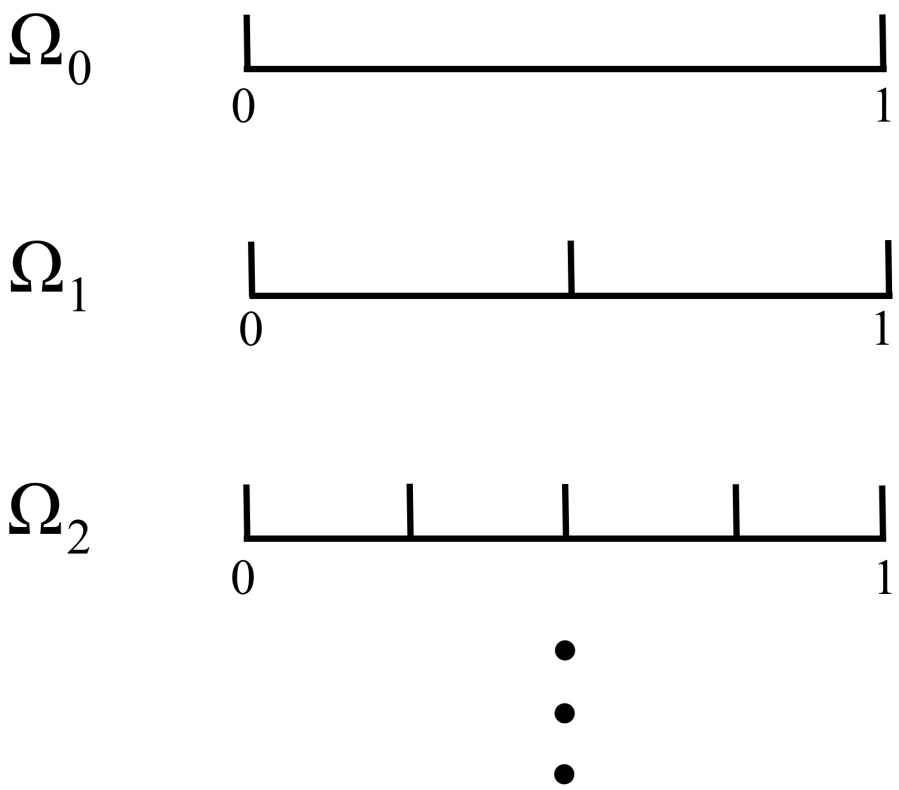
- **Sparse grid DG** method use multiwavelet (from MRA) and the DG weak form as building blocks.
- **Adaptive sparse grid DG** method perform thresholding based on hierarchical coefficients.
- For **nonlinear problems**, we developed new interpolatory multiwavelets.

Outline

The basics

Consider $\Omega = [0, 1]$ and define n-th level grid

$$\Omega_n = \{I_n^j = (2^{-n}j, 2^{-n}(j+1)], j = 0, \ldots, 2^n - 1\}$$



MRA

Conventional approximation space on the n-th level grid Ω_n

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

$$dim(V_n^k) = 2^n(k+1)$$

Nested structure

$$V_0^k \subset V_1^k \subset V_2^k \subset V_3^k \subset \cdots$$

W_n^k: orthogonal complement of V_{n-1}^k in V_n^k, for $n > 1$, represents the finer level details when the mesh is refined, satisfying

$$V_{n-1}^k \oplus W_n^k = V_n^k$$

$$W_n^k \perp V_{n-1}^k$$

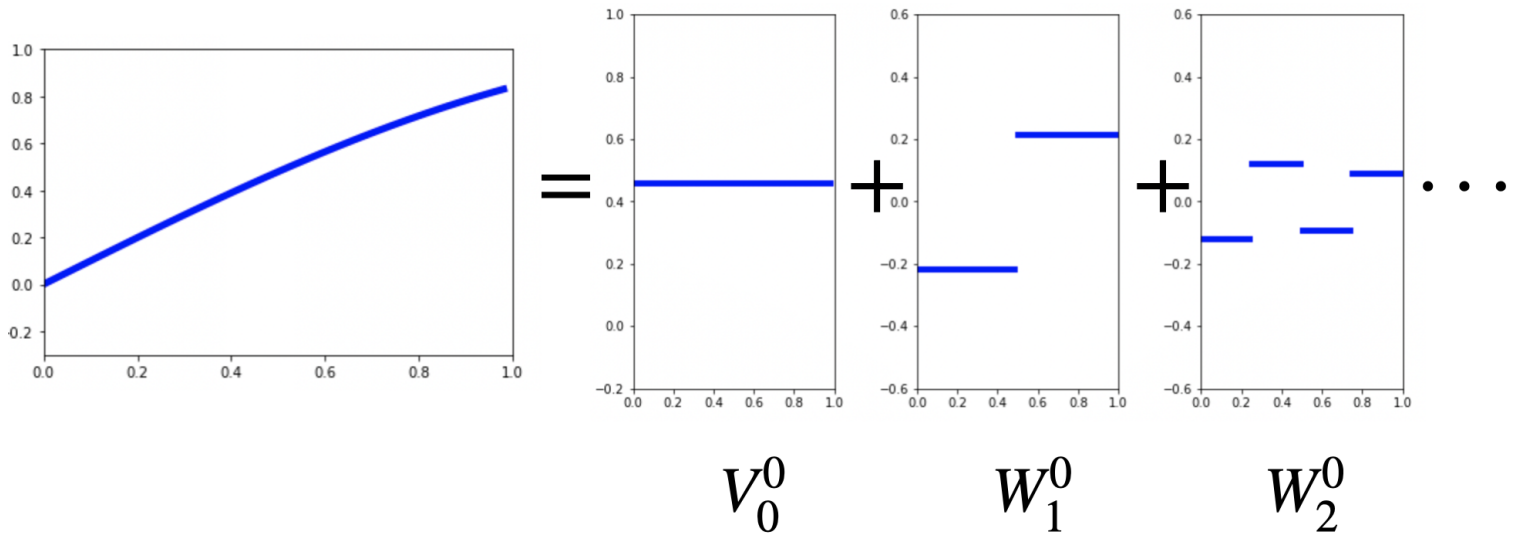
Let $W_0^k := V_0^k$, then

$$V_N^k = \bigoplus_{0 \le n \le N} W_n^k$$

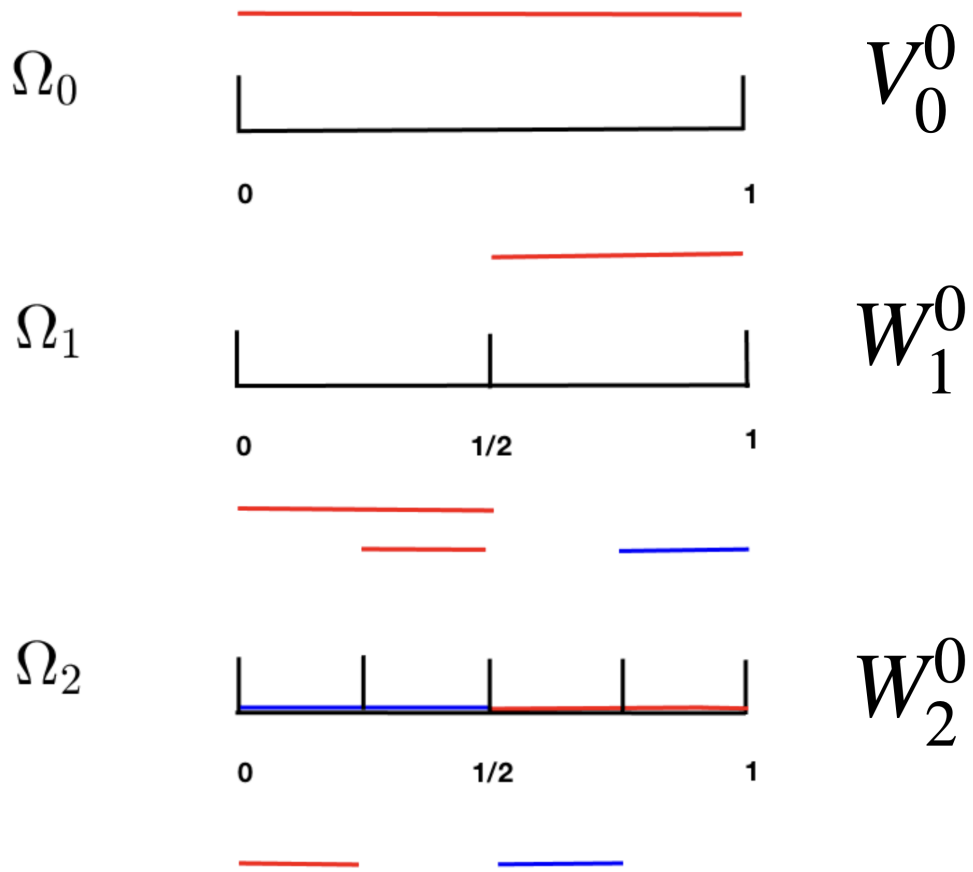
What does it mean?

Let P_n^k denotes the L^2 projection on to mesh level n, then

$$P_N^k f = \underbrace{P_0^k f}_{V_0^k} + \underbrace{(P_1^k - P_0^k)f}_{W_1^k} + \underbrace{(P_2^k - P_1^k)f}_{W_2^k} + \cdots + \underbrace{(P_N^k - P_{N-1}^k)f}_{W_N^k}$$

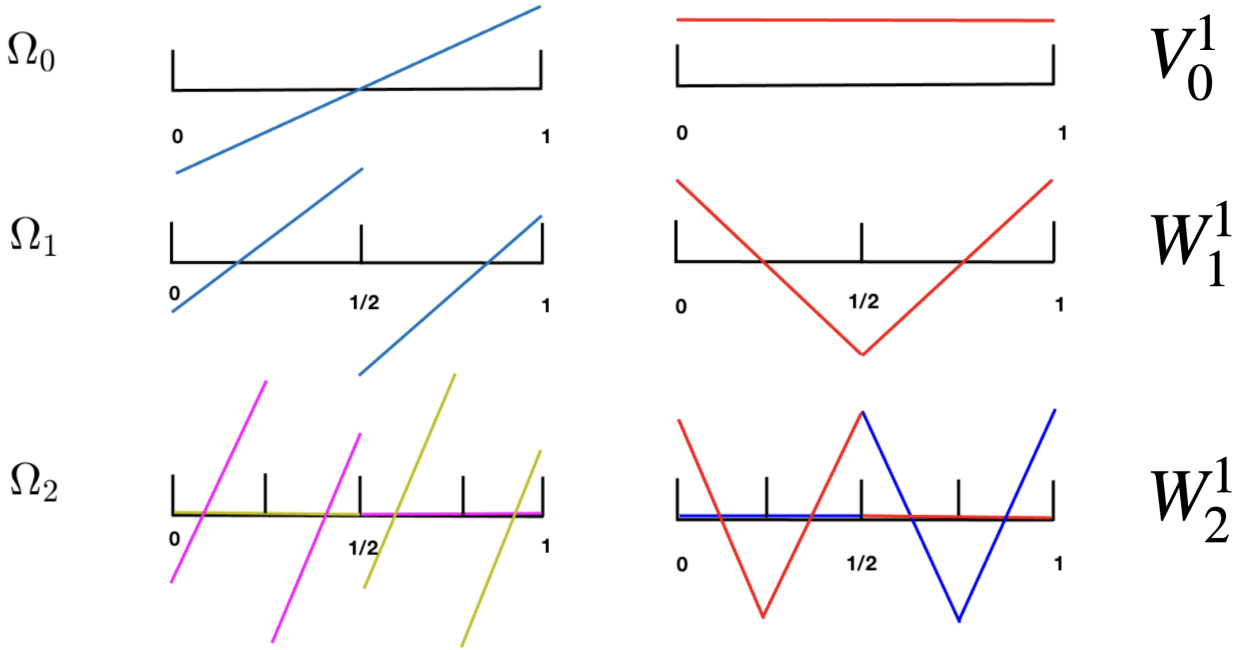


Bases on different levels for $k = 0$



Higher order

For $k = 1$



Arbitrary order k : we use L^2 orthogonal multiwavelets by Alpert (93).

Approximation space in multi-dimensions

Consider 2D case, $\mathbf{x} = (x_1, x_2) \in \Omega = [0,1]^2$ and multi-index $\mathbf{l} = (l_1, l_2) \in \mathbb{N}_0^2$
The standard rectangular grid $\Omega_{\mathbf{l}}$ with mesh size

$$h_{\mathbf{l}} := (2^{-l_1}, 2^{-l_2})$$
$$h := \min\{2^{-l_1}, 2^{-l_2}\}$$

For each $I_{\mathbf{l}}^{\mathbf{j}} = \{(x_1, x_2) : x_i \in (2^{-l_i} j_i, 2^{-l_i}(j_i + 1)]\}$, the traditional tensor-product polynomial space is

$$\mathbf{V}_{\mathbf{l}}^k = \{\mathbf{v} : \mathbf{v}(\mathbf{x}) \in P^k(I_{\mathbf{l}}^{\mathbf{j}}),\ \mathbf{0} \le \mathbf{j} \le 2^{\mathbf{l}} - \mathbf{1}\}$$

P^k denotes polynomial of degree at most k in each dimension.

Approximation space in multi-dimensions

Consider 2D case, $\mathbf{x} = (x_1, x_2) \in \Omega = [0,1]^2$ and multi-index $\mathbf{l} = (l_1, l_2) \in \mathbb{N}_0^2$
The standard rectangular grid $\Omega_\mathbf{l}$ with mesh size

$$h_\mathbf{l} := (2^{-l_1}, 2^{-l_2})$$
$$h := \min\{2^{-l_1}, 2^{-l_2}\}$$

For each $I_\mathbf{l}^\mathbf{j} = \{(x_1, x_2) : x_i \in (2^{-l_i} j_i, 2^{-l_i}(j_i + 1))]\}$, the traditional tensor-product polynomial space is

$$\mathbf{V}_\mathbf{l}^k = \{\mathbf{v} : \mathbf{v}(\mathbf{x}) \in P^k(I_\mathbf{l}^\mathbf{j}),\ \mathbf{0} \le \mathbf{j} \le 2^\mathbf{l} - \mathbf{1}\}$$

P^k denotes polynomial of degree at most k in each dimension. Uniform grid: $l_1 = l_2 = N$, $\mathbf{V}_\mathbf{l}^k = \mathbf{V}_N^k$, then

$$\mathbf{V}_N^k := V_{N,x_1}^k \times V_{N,x_2}^k = \bigoplus_{|\mathbf{l}|_\infty \le N} \mathbf{W}_\mathbf{l}^k$$

where

$$\mathbf{W}_\mathbf{l}^k := W_{l_1,x_1}^k \times W_{l_2,x_2}^k$$

The basis functions for $\mathbf{W}_\mathbf{l}^k$ can be defined by a tensor product

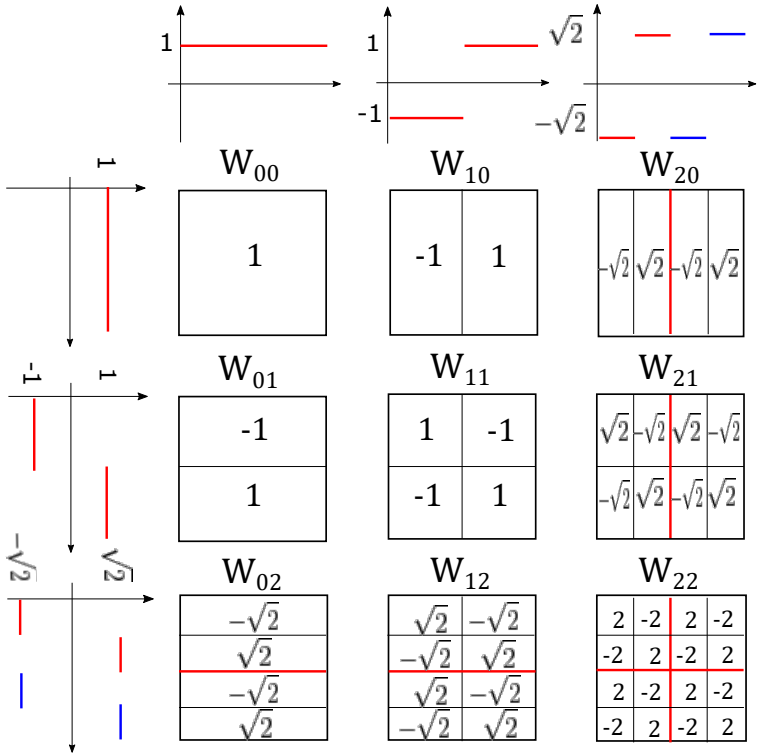
$$v_{\mathbf{i},\mathbf{l}}^\mathbf{j}(\mathbf{x}) := \prod_{t=1}^{2} v_{i_t,l_t}^{j_t}(x_t), \quad j_t = 0, \ldots, \max(0, 2^{l_t - 1} - 1), \quad i_t = 1, \ldots, k+1$$

Full grid approximation space

Full grid space:

$$\mathbf{V}_N^k = \bigoplus_{|\mathbf{l}|_\infty \le N} \mathbf{W}_\mathbf{l}^k$$

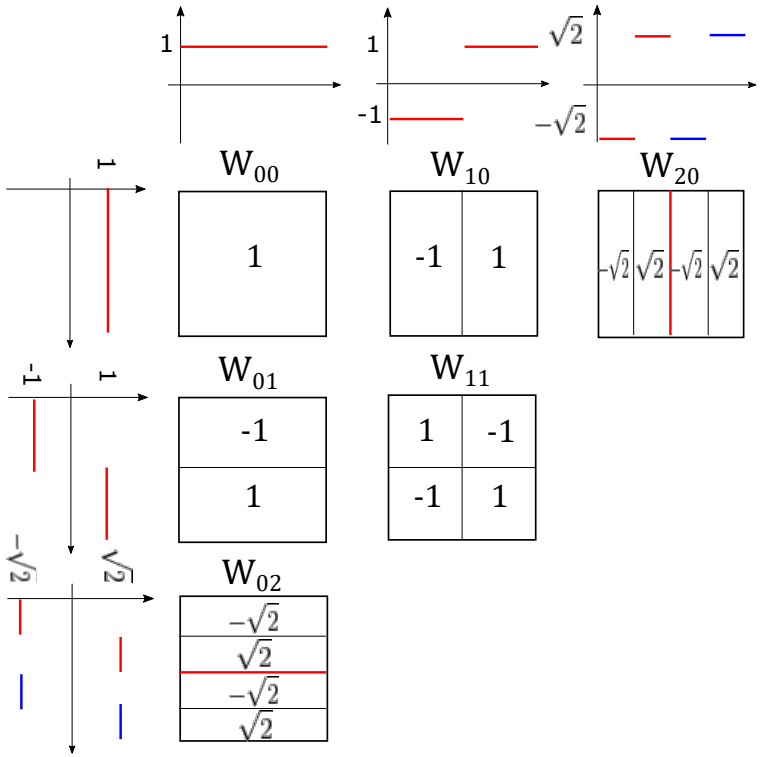
$d = 2,\ N = 2,\ k = 0$



$$dim(\mathbf{V}_N^k) = 2^{Nd}(k+1)^d \quad \text{or} \quad O(h^{-d})$$

Sparse grid approximation space

We consider the sparse grid space: $\hat{\mathbf{V}}_N^k := \bigoplus_{|\mathbf{l}|_1 \le N} \mathbf{W}_\mathbf{l}^k$



A viewpoint without using multiwavelet space: $\hat{\mathbf{V}}_N^k = \bigoplus_{|\mathbf{l}|_1 \le N} \mathbf{V}_\mathbf{l}^k$.

$$dim(\hat{\mathbf{V}}_N^k) = O(2^N N^{d-1}(k+1)^d) \quad \text{or} \quad O(h^{-1}|\log_2 h|^{d-1})$$

Sparse grid DG

Consider the linear transport equation with variable coefficient

$$\begin{cases} u_t + \nabla \cdot (\boldsymbol{\alpha}(\mathbf{x}, t)\, u) = 0, \quad \mathbf{x} \in \Omega = [0,1]^d, \\ u(0, \mathbf{x}) = u_0(\mathbf{x}), \end{cases} \tag{5}$$

The semi-discrete sparse grid DG [5] formulation for (5) is defined as follows: find $u_h \in \hat{\mathbf{V}}_N^k$, such that

$$\begin{aligned} \int_\Omega (u_h)_t\, v_h\, d\mathbf{x} &= \int_\Omega u_h \boldsymbol{\alpha} \cdot \nabla v_h\, d\mathbf{x} - \sum_{e \in \Gamma} \int_e \widehat{\boldsymbol{\alpha} u_h} \cdot [v_h]\, ds, \\ &\doteq A(u_h, v_h) \end{aligned} \tag{6}$$

for $\forall\, v_h \in \hat{\mathbf{V}}_N^k$, where $\widehat{\boldsymbol{\alpha} u_h}$ defined on the element interface denotes a monotone numerical flux.

[5] Guo, Cheng, SISC, 2016

Stability (constant coefficient case)

Theorem (Guo, Cheng, SISC, 2016)

The DG scheme (6) *for* (5) *is* L^2 *stable when* $\boldsymbol{\alpha}$ *is a constant vector, i.e.*

$$\frac{d}{dt}\int_{\Omega}(u_h)^2\,d\mathbf{x} = -\sum_{e\in\Gamma}\int_e \frac{|\boldsymbol{\alpha}\cdot\mathbf{n}|}{2}|[u_h]|^2 ds \leq 0. \tag{7}$$

Error estimate (constant coefficient case)

Theorem (Guo, Cheng, SISC, 2016)

Let u be the exact solution, and u_h be the numerical solution to the semi-discrete scheme (6) with numerical initial condition $u_h(0) = \mathbf{P}u_0$. For $k \geq 1$, $u_0 \in \mathcal{H}^{p+1}(\Omega)$, $1 \leq q \leq \min\{p, k\}$, $N \geq 1$, $d \geq 2$, we have for all $t \geq 0$,

$$\|u_h - u\|_{L^2(\Omega_N)} \leq$$
$$\left(2\sqrt{C_d||\boldsymbol{\alpha}||_2 t}\, C_\star(k,q,d,N) + (\bar{\bar{c}}_{k,0,q} + B_0(k,q,d)\kappa_0(k,q,N)^d)2^{-N/2}\right) 2^{-N(q+1/2)}|u_0|_{\mathcal{H}^{q+1}(\Omega)},$$

where C_d is a generic constant with dependence only on d, $C_\star(k,q,d,N) = \max_{s=0,1}\left(\bar{\bar{c}}_{k,s,q} + B_s(k,q,d)\kappa_s(k,q,N)^d\right)$. The constants $\bar{\bar{c}}_{k,s,q}$, $B_s(k,q,d)$, $\kappa_s(k,q,N)$ are defined in L^2 projection error estimates.

Convergence rate $O((\log h)^d h^{k+1/2})$.

Linear advection: sparse grid DG

We consider the following linear advection problem

$$\begin{cases} u_t + \sum_{m=1}^{d} u_{x_m} = 0, \quad \mathbf{x} \in [0,1]^d, \\ u(0,\mathbf{x}) = \sin\left(2\pi \sum_{m=1}^{d} x_m\right), \end{cases} \tag{8}$$

subject to periodic boundary conditions.
In the simulation, we compute the numerical solutions up to two periods in time, meaning that we let final time $T = 1$ for $d = 2$, $T = 2/3$ for $d = 3$, and $T = 0.5$ for $d = 4$.

Error and DOF

N	h_N	DOF	L^2 error	order	DOF	L^2 error	order
		$k=1,\ d=3$			$k=1,\ d=4$		
4	1/16	832	3.72E-01	–	3072	4.99E-01	–
5	1/32	2176	1.19E-01	1.64	8832	2.40E-01	1.06
6	1/64	5504	2.96E-02	2.01	24320	9.84E-02	1.28
7	1/128	13568	8.85E-03	1.74	64768	3.21E-02	1.62
		$k=2,\ d=3$			$k=2,\ d=4$		
4	1/16	2808	1.10E-02	–	15552	2.80E-02	–
5	1/32	7344	1.79E-03	2.63	44712	5.82E-03	2.27
6	1/64	18576	3.97E-04	2.17	123120	1.37E-03	2.09
7	1/128	45792	5.14E-05	2.95	327888	2.58E-04	2.41

Table: L^2 errors and orders of accuracy, DOF.

FG: 56Million

FG: 21Billion

Outline

Adaptivity

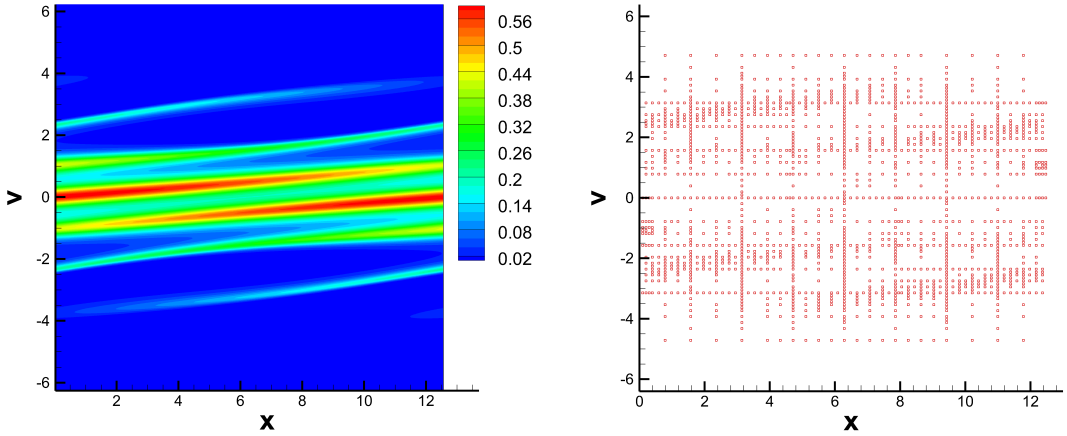
- Sparse grid has poor resolution when function is not smooth.
- We developed adaptive sparse grid DG method (Guo, Cheng, SISC, 2017) to address this issue.
- The idea is to threshold based on the hierarchical coefficients, like MRA for image processing.

Predict → Refine → Evolve → Coarsen

- Note: when the solution is regular, adaptive sparse grid will return to standard sparse grid method, retaining its advantage for high dimensional problems.

Numerical example

Vlasov-Poisson/Vlasov-Maxwell up to 4D. Example: Landau damping $t = 10.$[6]



(a) Solution (b) Active element

[6] Guo, Cheng, SISC, 2017

Nonlinear problems

- Example: nonlinear source term $f(u)$ requires evaluating terms

$$\int_\Omega f(u_h)v_h dx = \sum_K \int_K f(u_h)v_h dx,$$

  where u_h is represented by multiwavelet basis functions.
- We cannot afford to sum up on all elementary cells K as this requires $O(h^{-d})$ operations.
- The idea is to switch to nodal basis and evaluate

$$\int_\Omega \mathcal{I}f(u_h)v_h dx.$$

- Next, we will demonstrate the construction of $\mathcal{I}$.

1D: nested points

Consider the domain $I = [0, 1]$, we use the same notation.In addition, we define $k + 1$ distinct points on each cell

$$x_{i,n}^j = 2^{-n} j + 2^{-n} \alpha_i \tag{9}$$

with $\alpha_i \in [0, 1]$, $i = 1, \ldots, k + 1$.
In particular, the collection of those points $X_n^k = \{x_{i,n}^j\}$ is called *nested points*, if

$$X_0^k \subset X_1^k \subset X_2^k \subset \cdots . \tag{10}$$

1D

Since $\{X_n^k\}$ are nested, the points can be rearranged in such a way that

$$X_n^k = X_0^k \cup \widetilde{X}_1^k \cup \cdots \cup \widetilde{X}_n^k, \quad \text{with } \widetilde{X}_n^k = X_n^k / X_{n-1}^k. \tag{11}$$

Moreover, we can now define the subspace $\widetilde{W}_n^k$, $n \geq 1$, as the complement of V_{n-1}^k in V_n^k, in which the piecewise polynomials vanish at all points in X_{n-1}^k,

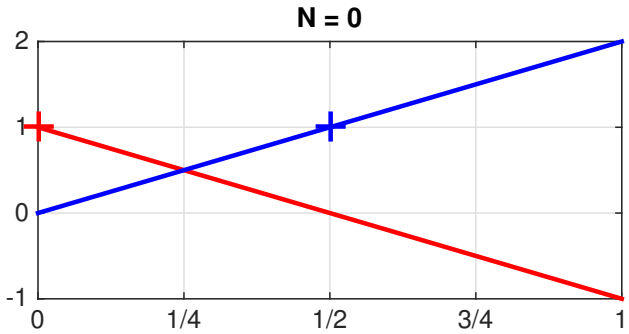
$$V_n^k = V_{n-1}^k \oplus \widetilde{W}_n^k. \tag{12}$$

This corresponds to

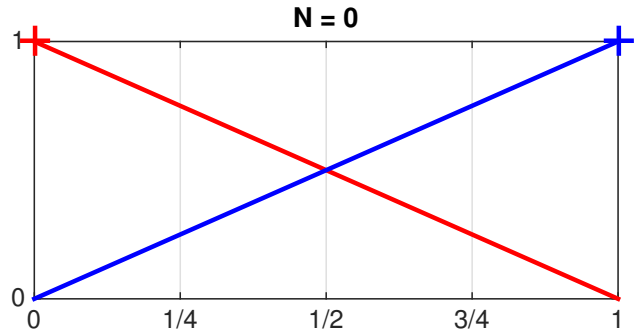
$$I_N^k f = \underbrace{I_0^k f}_{V_0^k} + \underbrace{(I_1^k - I_0^k) f}_{\widetilde{W}_1^k} + \underbrace{(I_2^k - I_1^k) f}_{\widetilde{W}_2^k} + \cdots + \underbrace{(I_N^k - I_{N-1}^k) f}_{\widetilde{W}_N^k}$$

[6] Tao, Jiang, Cheng JCP 2021

1D-Example



(c) P^1: $x_0 = 0$, $x_1 = 1/2$



(d) P^1: $x_0 = 0$, $x_1 = 1$

Figure: Interpolation points and multiwavelets: P^1.

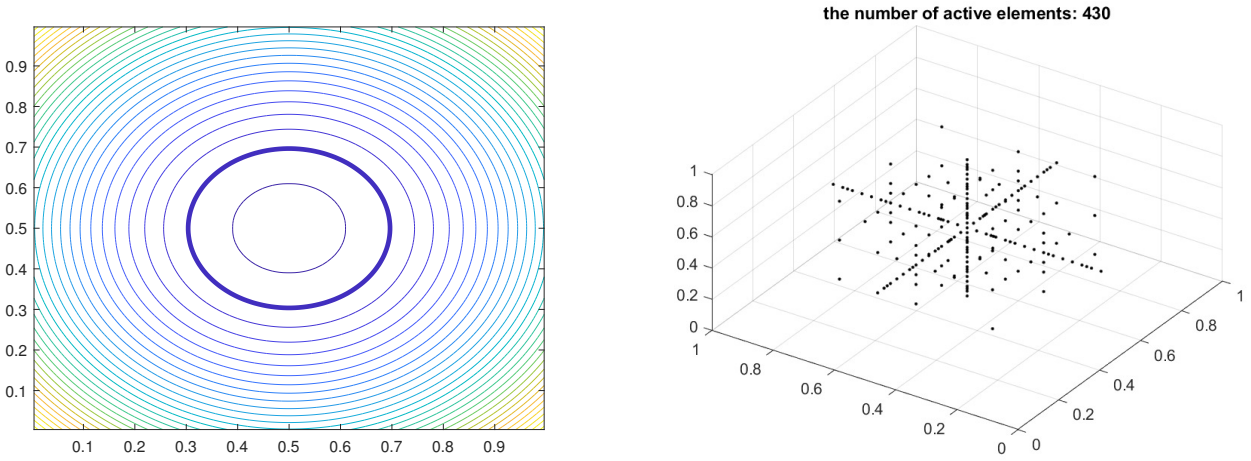
[6] Tao, Jiang, Cheng JCP 2021

Summary

- Similarly, the approach extends to Hermite interpolation. One can convert between point (derivative) values to multiwavelet coefficients using fast wavelet transform.
- For multi-D, the approach works for both sparse grid and adaptive sparse grid.
- Incorporating this into numerical schemes requires a bit more than fast wavelet transform. Fast matrix-vector product Shen, Yu (10, 12) is needed. We show in Huang, Cheng (20), Huang, Guo, Cheng (22) how the algorithm works.

Numerical example: HJ/HJB equations (with LDG solver)

3D Eikonal



(a) (b)

Figure: (a) 2D cut. (b) Active elements. Error tolerance ϵ=10^{-7}.

[6] Guo, Huang, Tao, Cheng JCP 2021

Outline

Conclusions and Outlook

- To design Sparse grid DG solvers, we need tools from *numerical PDEs, signal processing and approximation theory.* I find this journey very inspiring and enriching.
- In high dimensional scientific computing, another main compression technique is the **low rank method**. We plan to investigate low rank tensor methods for solving time dependent PDE models. This will draw on knowledge from numerical PDEs and linear algebra.
- There is also whole arsenal of tools in 'model order reduction'. We have developed some ROMs for kinetic equations by reduced basis methods and machine learning. We are interested in exploring the data driven modeling aspects, perhaps in combination with the high dimensional solvers mentioned before.

The END!

Thank You!