

Efficient computation of Jacobian matrices for entropy stable summation-by-parts schemes

Jesse Chan, Christina Taylor

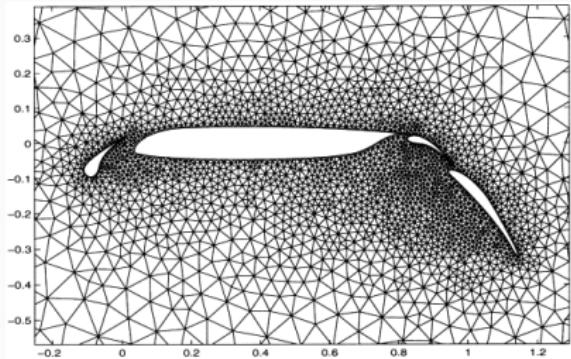
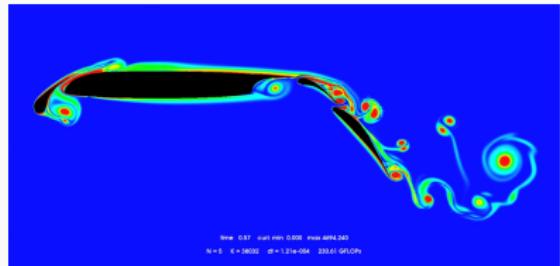
SIAM TX-LA Meeting

October 17, 2020

Department of Computational and Applied Mathematics

High order finite element methods for hyperbolic PDEs

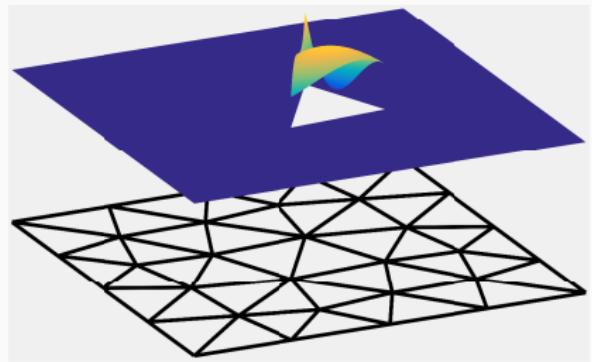
- Aerodynamics applications:
acoustics, vorticular flows,
turbulence, shocks.
- Goal: high **accuracy** simulations
on **unstructured meshes**.
- Discontinuous Galerkin (DG)
methods: geometric flexibility,
high order accuracy.



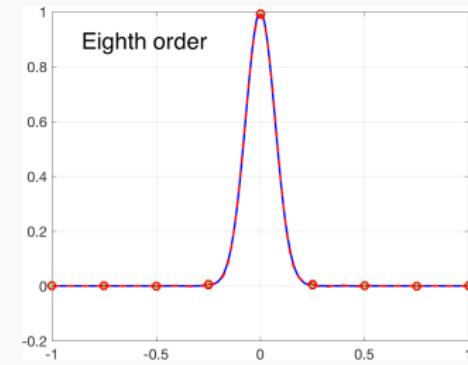
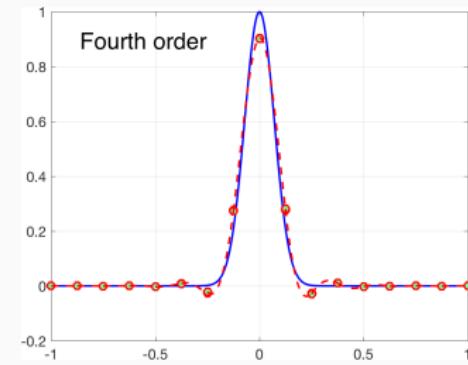
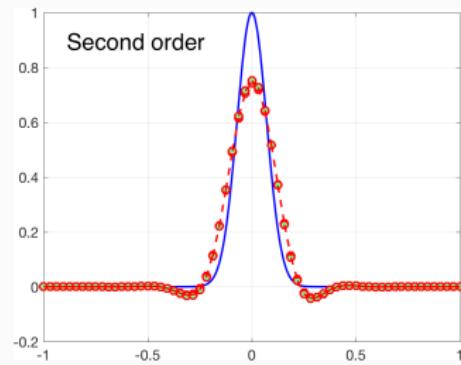
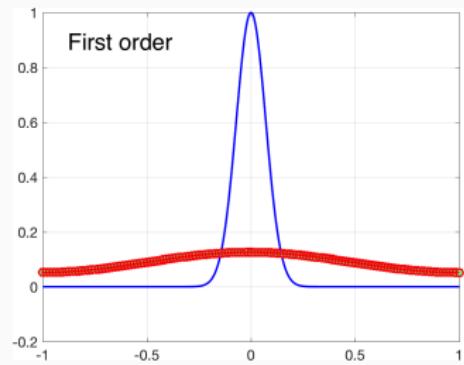
Mesh from Slawig 2001.

High order finite element methods for hyperbolic PDEs

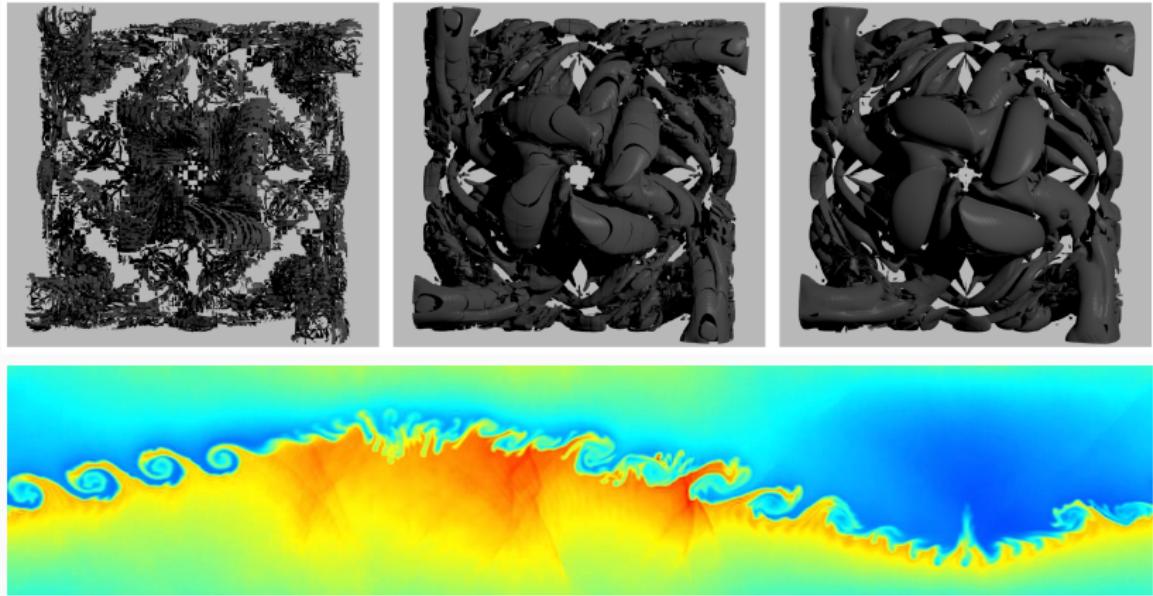
- Aerodynamics applications:
acoustics, vorticular flows,
turbulence, shocks.
- Goal: high **accuracy** simulations
on **unstructured meshes**.
- Discontinuous Galerkin (DG)
methods: geometric flexibility,
high order accuracy.



Why high order accuracy?

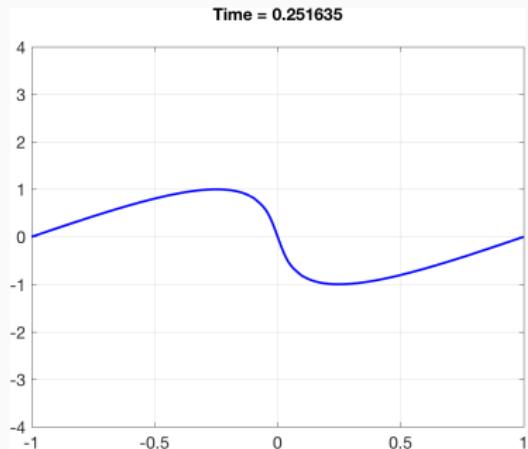
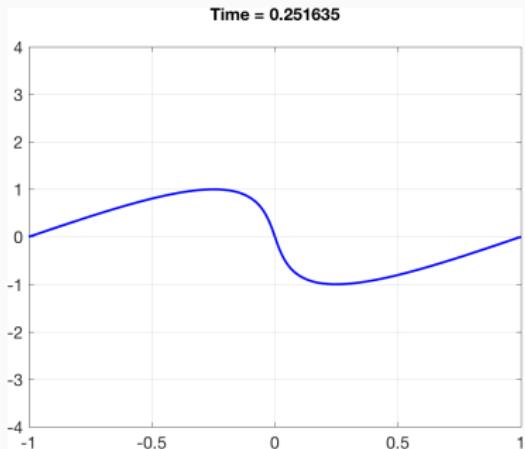


Why high order accuracy?



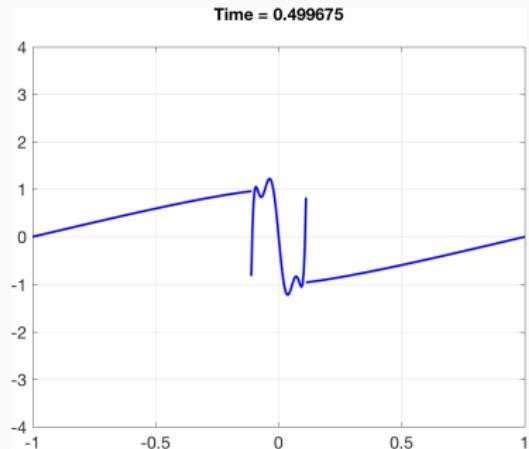
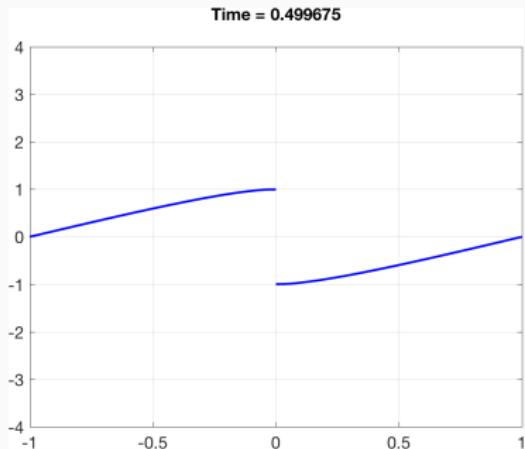
2nd, 4th, and 16th order Taylor-Green (top), 8th order Kelvin-Helmholtz (bottom). Vorticlar structures and acoustic waves are both sensitive to numerical dissipation. Results from Beck and Gassner (2013) and Per-Olof Persson's website.

Why *not* high order DG methods?



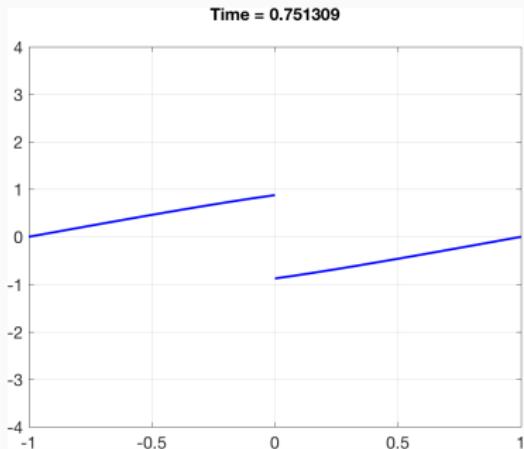
High order methods blow up for under-resolved solutions of **nonlinear conservation laws** (e.g., shocks and turbulence).

Why *not* high order DG methods?

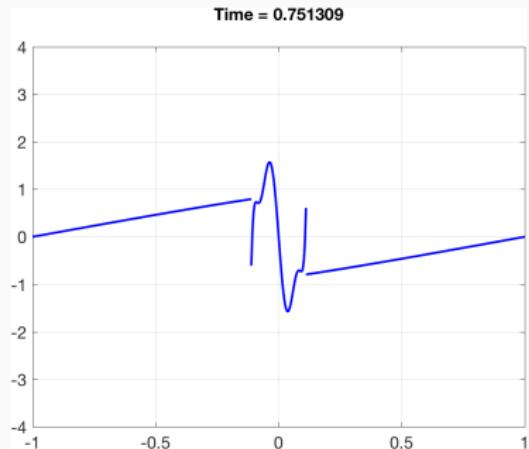


High order methods blow up for under-resolved solutions of **nonlinear conservation laws** (e.g., shocks and turbulence).

Why *not* high order DG methods?



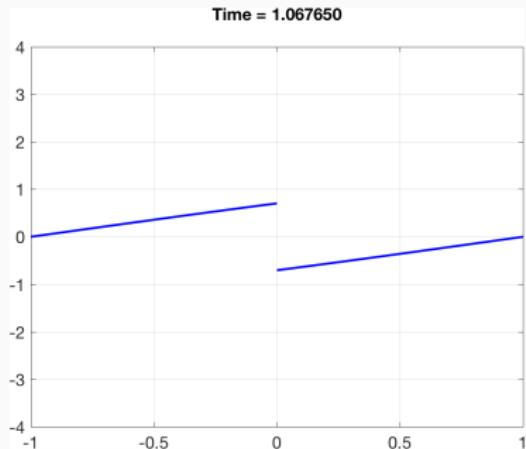
(a) Exact solution



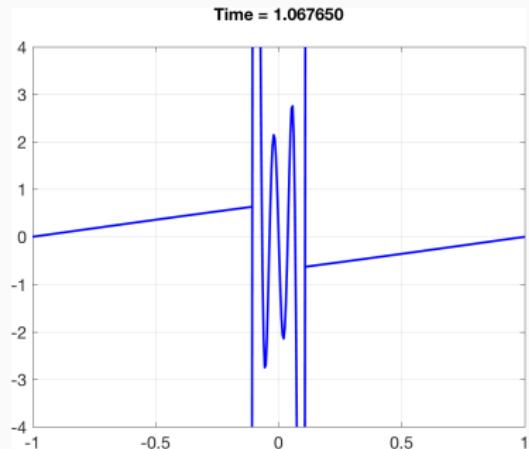
(b) 8th order DG

High order methods blow up for under-resolved solutions of **nonlinear conservation laws** (e.g., shocks and turbulence).

Why *not* high order DG methods?



(a) Exact solution



(b) 8th order DG

High order methods blow up for under-resolved solutions of **nonlinear conservation laws** (e.g., shocks and turbulence).

Continuous stability $\not\Rightarrow$ discrete stability

- Entropy stability: generalizes energy stability to **nonlinear** conservation laws (shallow water, compressible Euler + Navier-Stokes).

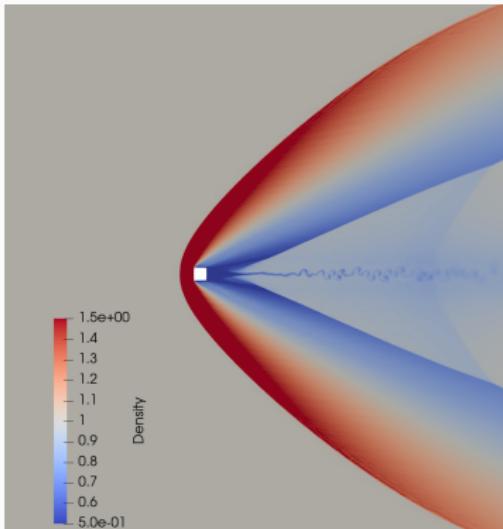
$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{u})}{\partial x} = 0.$$

- Continuous entropy inequality: given a convex **entropy** function $S(\mathbf{u})$ and “entropy potential” $\psi(\mathbf{u})$, test with $\mathbf{v}(\mathbf{u})$

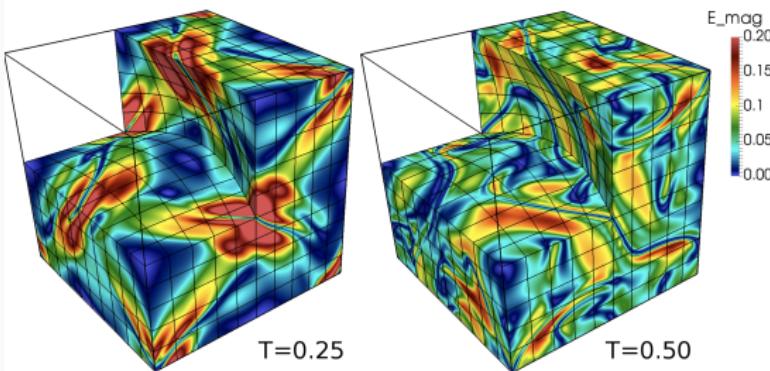
$$\int_{\Omega} \mathbf{v}^T \left(\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{u})}{\partial x} \right) = 0, \quad \boxed{\mathbf{v}(\mathbf{u}) = \frac{\partial S}{\partial \mathbf{u}}} \\ \Rightarrow \int_{\Omega} \frac{\partial S(\mathbf{u})}{\partial t} + (\mathbf{v}^T \mathbf{f}(\mathbf{u}) - \psi(\mathbf{u}))|_{-1}^1 \leq 0.$$

- Proof of entropy inequality relies on **chain rule**, integration by parts.

Entropy stable methods recover a discrete entropy inequality.



(a) Compressible Navier-Stokes
($\text{Ma}=1.5$, $\text{Re}=10000$, degree 3)



(b) Resistive MHD (3D Orszag-Tang)

No limiters, filters, or artificial viscosity beyond DG “upwinded” fluxes.

Talk outline

- 1 Entropy stable nodal summation-by-parts (SBP) schemes
- 2 Jacobian matrices for “flux differencing” formulations
- 3 Numerical experiments

Talk outline

- 1 Entropy stable nodal summation-by-parts (SBP) schemes
- 2 Jacobian matrices for “flux differencing” formulations
- 3 Numerical experiments

Ingredients for entropy stable methods

- Summation-by-parts (SBP) differentiation matrices \mathbf{Q}_i with boundary matrix \mathbf{B}_i

$$\mathbf{Q}_i + \mathbf{Q}_i^T = \mathbf{B}_i.$$

- Tadmor-type entropy conservative flux

$$\mathbf{f}_S(\mathbf{u}, \mathbf{u}) = \mathbf{f}(\mathbf{u}), \quad (\text{consistency})$$

$$\mathbf{f}_S(\mathbf{u}, \mathbf{v}) = \mathbf{f}_S(\mathbf{v}, \mathbf{u}), \quad (\text{symmetry})$$

$$(\mathbf{v}_L - \mathbf{v}_R)^T \mathbf{f}_S(\mathbf{u}_L, \mathbf{u}_R) = \psi_L - \psi_R, \quad (\text{conservation}).$$

Flux differencing formulations

- Main idea: “flux differencing” discretization

$$\mathbf{M} \frac{d\mathbf{u}}{dt} + \sum_{i=1}^d \left((\mathbf{Q}_i - \mathbf{Q}_i^T) \circ \mathbf{F}_i \right) \mathbf{1} + \mathbf{B}_i \mathbf{f}^* = \mathbf{0}.$$

$$(\mathbf{F}_i)_{jk} = f_{i,S}(\mathbf{u}_j, \mathbf{u}_k)$$

- Can prove discrete entropy conservation

$$\mathbf{1}^T \mathbf{M} \frac{dS(\mathbf{u})}{dt} = 0, \quad \text{for appropriate BCs}$$

- Add numerical or physical dissipation for discrete entropy inequality

$$\mathbf{1}^T \mathbf{M} \frac{dS(\mathbf{u})}{dt} + \underbrace{\mathbf{v}^T (\mathbf{K} \circ \mathbf{D}) \mathbf{1}}_{\geq 0} = 0$$

Example 1: entropy conservative finite volume methods

- Usual formulation

$$\frac{d\mathbf{u}_i}{dt} + \frac{\mathbf{f}_S(\mathbf{u}_{i+1}, \mathbf{u}_i) - \mathbf{f}_S(\mathbf{u}_i, \mathbf{u}_{i-1})}{h} = \mathbf{0}$$

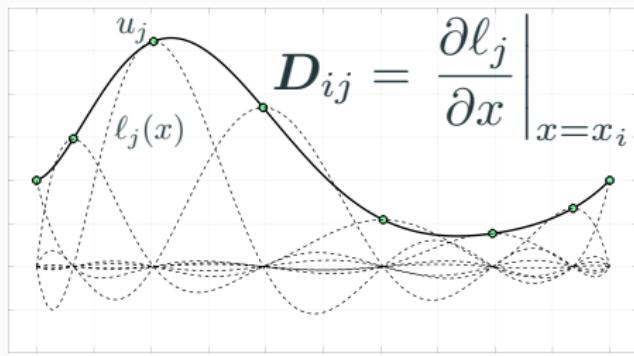
- Mass matrix $\mathbf{M} = h\mathbf{I}$, differentiation matrix \mathbf{Q}

$$\mathbf{Q} = \frac{1}{2} \begin{bmatrix} -1 & 1 & & \\ -1 & & \ddots & \\ & \ddots & & 1 \\ & & -1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}.$$

- Let $\mathbf{f}^* = [\mathbf{f}_L, 0, \dots, 0, \mathbf{f}_R]$ impose BCs. Equivalent to

$$\mathbf{M} \frac{d\mathbf{u}}{dt} + ((\mathbf{Q} - \mathbf{Q}^T) \circ \mathbf{F}) \mathbf{1} + \mathbf{B} \mathbf{f}^* = \mathbf{0}.$$

Example 2: high order spectral collocation



$$D_{ij} = \frac{\partial \ell_j}{\partial x} \Big|_{x=x_i}$$

$$\mathbf{B} = \begin{bmatrix} -1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

- SBP operators at Lobatto quadrature points: lumped mass matrix \mathbf{M} , weak differentiation matrix $\mathbf{Q} = \mathbf{MD}$.
- Entropy conservative formulation: same algebraic structure

$$\mathbf{M} \frac{d\mathbf{u}}{dt} + ((\mathbf{Q} - \mathbf{Q}^T) \circ \mathbf{F}) \mathbf{1} + \mathbf{Bf}^* = \mathbf{0}.$$

Example 3: high order modal DG formulations

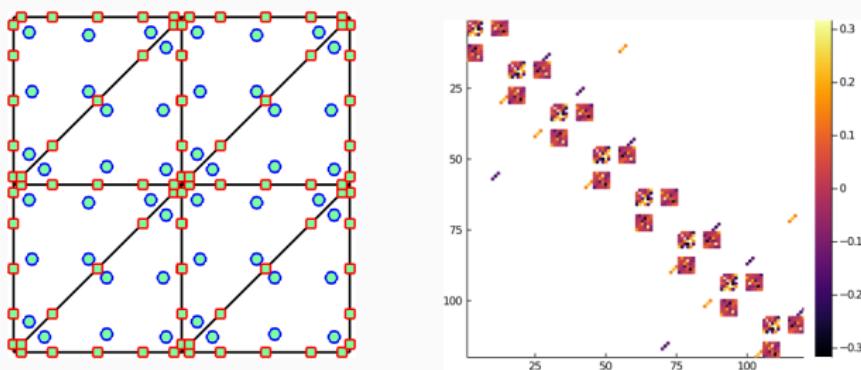


Figure: Global SBP-DG matrix for $N = 2$ on a periodic mesh.

- $\mathbf{V}_h, \mathbf{V}_f$ matrices interpolate basis coefficients to quadrature.
- Formulation with global SBP matrix \mathbf{Q}_i .

$$\mathbf{M} \frac{d\mathbf{u}}{dt} + \sum_{i=1}^d \mathbf{V}_h^T ((\mathbf{Q}_i - \mathbf{Q}_i^T) \circ \mathbf{F}_i) \mathbf{1} + \mathbf{V}_f^T \mathbf{B} \mathbf{f}^* = \mathbf{0}.$$

Example 3: high order modal DG formulations, cont.

- “Modal” formulation uses L^2 projection of entropy variables $\mathbf{v}(\mathbf{u}_h)$

$$\Pi_N \mathbf{v}(\mathbf{u}_h), \quad \Pi_N = L^2 \text{ projection onto degree } N \text{ polynomials}$$

- Flux matrix \mathbf{F}_i must be computed in terms of the *entropy-projected conservative variables* $\tilde{\mathbf{u}}$

$$(\mathbf{F}_i)_{jk} = f_{i,S}(\tilde{\mathbf{u}}_j, \tilde{\mathbf{u}}_k)$$

$$\tilde{\mathbf{u}} \approx \mathbf{u}(\Pi_N \mathbf{v}(\mathbf{u}_h)).$$

- Enables near-arbitrary combinations of basis and quadrature.

Example 3: high order modal DG formulations, cont.

- “Modal” formulation uses L^2 projection of entropy variables $\mathbf{v}(\mathbf{u}_h)$

$$\Pi_N \mathbf{v}(\mathbf{u}_h), \quad \Pi_N = L^2 \text{ projection onto degree } N \text{ polynomials}$$

- Flux matrix \mathbf{F}_i must be computed in terms of the *entropy-projected conservative variables* $\tilde{\mathbf{u}}$

$$(\mathbf{F}_i)_{jk} = \mathbf{f}_{i,S}(\tilde{\mathbf{u}}_j, \tilde{\mathbf{u}}_k)$$
$$\tilde{\mathbf{u}} \approx \mathbf{u}(\Pi_N \mathbf{v}(\mathbf{u}_h)).$$

- Enables near-arbitrary combinations of basis and quadrature.

Example 3: high order modal DG formulations, cont.

- “Modal” formulation uses L^2 projection of entropy variables $\mathbf{v}(\mathbf{u}_h)$

$$\Pi_N \mathbf{v}(\mathbf{u}_h), \quad \Pi_N = L^2 \text{ projection onto degree } N \text{ polynomials}$$

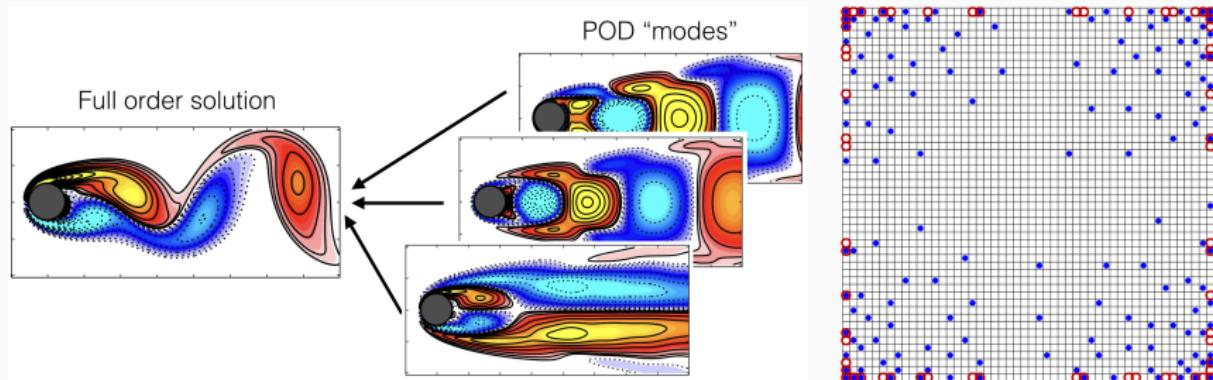
- Flux matrix \mathbf{F}_i must be computed in terms of the *entropy-projected conservative variables* $\tilde{\mathbf{u}}$

$$(\mathbf{F}_i)_{jk} = \mathbf{f}_{i,S}(\tilde{\mathbf{u}}_j, \tilde{\mathbf{u}}_k)$$

$$\tilde{\mathbf{u}} \approx \mathbf{u}(\Pi_N \mathbf{v}(\mathbf{u}_h)).$$

- Enables near-arbitrary combinations of basis and quadrature.

Example 4: entropy stable reduced order modeling

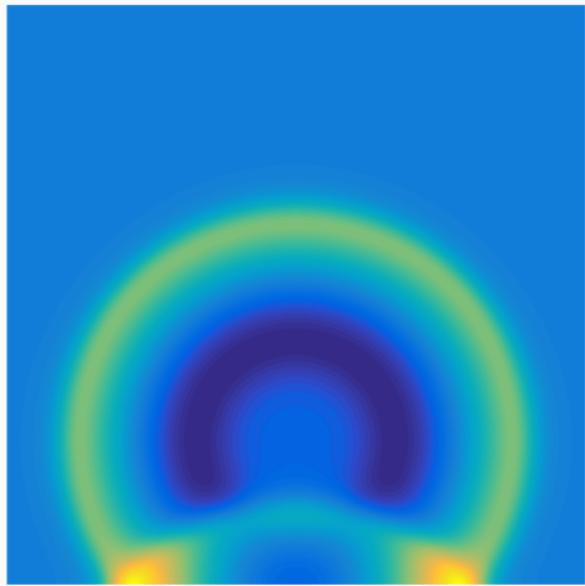


(a) Figure adapted from Brunton, Proctor, Kutz (2016)

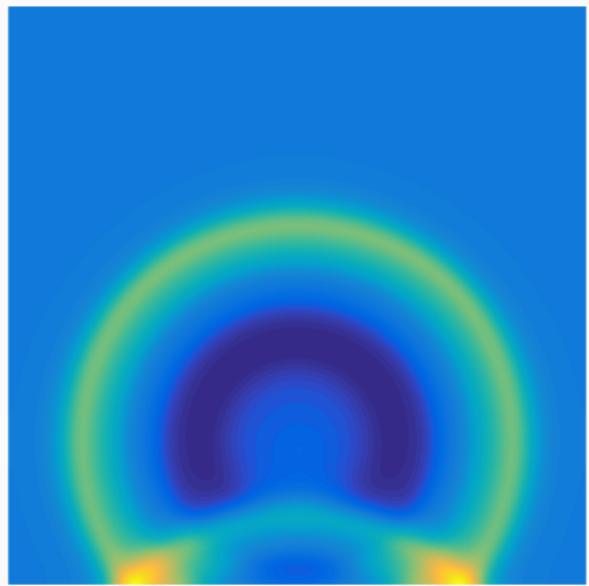
(b) Hyper-reduction

- Principal orthogonal decomposition (POD) \iff basis functions
- Sampling/weighting hyper-reduction \iff quadrature
- Treat as single-element modal DG scheme.

Example 4: entropy stable reduced order modeling, cont.



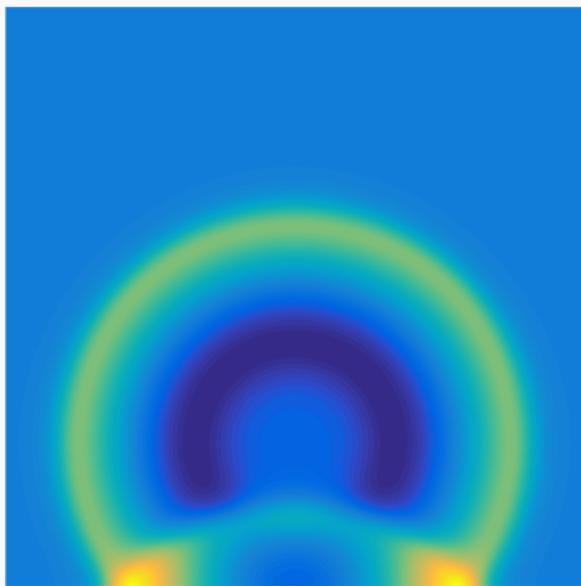
(a) Density, full order model



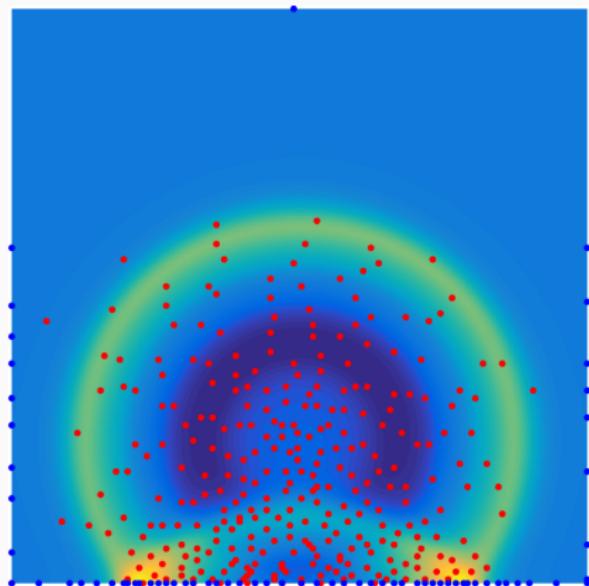
(b) Reduced order model

Full order model with 10000 points, ROM with 25 modes, 306 points.

Example 4: entropy stable reduced order modeling, cont.



(a) Density, full order model



(b) ROM w/reduced quad. points

Full order model with 10000 points, ROM with 25 modes, 306 points.

Talk outline

- 1 Entropy stable nodal summation-by-parts (SBP) schemes
- 2 Jacobian matrices for “flux differencing” formulations
- 3 Numerical experiments

Current methods for computing Jacobian matrices

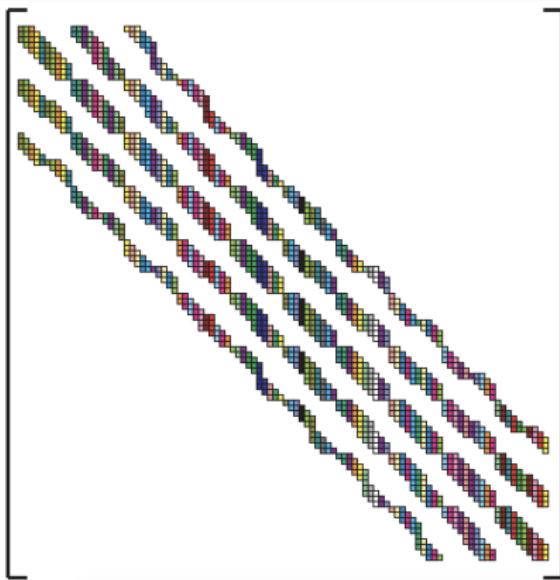
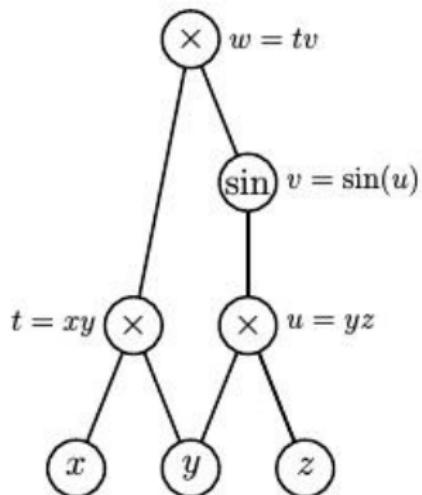


Figure from Gebremedhin, Manne, Pothen (2005),
*What color is your Jacobian? Graph coloring for
computing derivatives.*

- Matrix-free approach if only computing Jacobian-vector products
- Compute entries using finite differences or automatic differentiation (AD)
- Graph coloring reduces function evaluations and AD costs, but only for **sparse** matrices

$$\mathbf{J}(\mathbf{u})\Delta\mathbf{u} \approx \frac{\mathbf{f}(\mathbf{u} + \Delta\mathbf{u}) - \mathbf{f}(\mathbf{u})}{\|\Delta\mathbf{u}\|}.$$

Cost of computing dense Jacobian blocks



- Graph coloring AD expensive for dense matrix blocks (e.g., high order DG methods, reduced order models).
- Problem: cost of AD scales with size of input and output dimension.
- Can reduce costs by only applying AD only to nonlinear flux $f_S(\mathbf{u}_L, \mathbf{u}_R)$.

Image from Austin (2017), *How to Differentiate with a Computer*.

Jacobian matrices for flux differencing

Hadamard product structure of flux differencing yields simple Jacobians.

Theorem

Assume $\mathbf{Q} = \pm \mathbf{Q}^T$. Consider a scalar "collocation" discretization

$$\mathbf{r}(\mathbf{u}) = (\mathbf{Q} \circ \mathbf{F}) \mathbf{1}, \quad \mathbf{F}_{ij} = f_S(\mathbf{u}_i, \mathbf{u}_j).$$

The Jacobian matrix is then

$$\frac{d\mathbf{r}}{d\mathbf{u}} = (\mathbf{Q} \circ \partial\mathbf{F}_R) \pm \text{diag} \left(\mathbf{1}^T (\mathbf{Q} \circ \partial\mathbf{F}_R) \right),$$

$$(\partial\mathbf{F}_R)_{ij} = \left. \frac{\partial f_S(u_L, u_R)}{\partial u_R} \right|_{\mathbf{u}_i, \mathbf{u}_j}.$$

Observations about flux differencing Jacobian formulas

Separates "template" discretization matrix \mathbf{Q} and flux contributions.

$$\frac{d\mathbf{r}}{d\mathbf{u}} = (\mathbf{Q} \circ \partial\mathbf{F}_R) \pm \text{diag}(\mathbf{1}^T (\mathbf{Q} \circ \partial\mathbf{F}_R)),$$

$$(\partial\mathbf{F}_R)_{ij} = \left. \frac{\partial f_S(u_L, u_R)}{\partial u_R} \right|_{\mathbf{u}_i, \mathbf{u}_j}.$$

Option 1: compute derivatives $\frac{\partial f_S(u_L, u_R)}{\partial u_R}$ analytically

$$f_S(u_L, u_R) = \frac{1}{6} (u_L^2 + u_L u_R + u_R^2)$$

$$\frac{\partial f_S(u_L, u_R)}{\partial u_R} = \frac{1}{6} (u_L + 2u_R).$$

Observations about flux differencing Jacobian formulas

Separates "template" discretization matrix \mathbf{Q} and flux contributions.

$$\frac{d\mathbf{r}}{d\mathbf{u}} = (\mathbf{Q} \circ \partial\mathbf{F}_R) \pm \text{diag}(\mathbf{1}^T (\mathbf{Q} \circ \partial\mathbf{F}_R)),$$

$$(\partial\mathbf{F}_R)_{ij} = \left. \frac{\partial f_S(u_L, u_R)}{\partial u_R} \right|_{\mathbf{u}_i, \mathbf{u}_j}.$$

Option 2: use AD for $\frac{\partial f_S(u_L, u_R)}{\partial u_R}$. Efficient: $O(1)$ inputs/outputs. In Julia:

```
using ForwardDiff
```

```
f(uL, uR) = (1/6) * (uL^2 + uL*uR + uR^2)
```

```
dF(uL, uR) = ForwardDiff.derivative(uR->f(uL, uR), uR)
```

Fluxes can be complicated to differentiate analytically

- Entropy conservative fluxes for 1D compressible Euler

$$f_S^1(\mathbf{u}_L, \mathbf{u}_R) = \{\{\rho\}\}^{\log} \{\{u\}\}$$

$$f_S^2(\mathbf{u}_L, \mathbf{u}_R) = \frac{\{\{\rho\}\}}{2 \{\{\beta\}\}} + \{\{u\}\} f_S^1$$

$$f_S^3(\mathbf{u}_L, \mathbf{u}_R) = f_S^1 \left(\frac{1}{2(\gamma - 1) \{\{\beta\}\}^{\log}} - \frac{1}{2} \{\{u^2\}\} \right) + \{\{u\}\} f_S^2,$$

- Fluxes involve logarithmic mean $\{\{u\}\}^{\log} = \frac{u_L - u_R}{\log u_L - \log u_R}$ and "inverse temperature" $\beta = \frac{\rho}{2p}$.
- Specialized evaluation of $\{\{u\}\}^{\log}$ using γ -expansions.

Extensions: systems of nonlinear conservation laws

- Assume n fields; nonlinear term is now

$$\mathbf{r}(\mathbf{u}) = ((\mathbf{I}_n \otimes \mathbf{Q}) \circ \mathbf{F}) \mathbf{1} = \begin{bmatrix} (\mathbf{Q} \circ \mathbf{F}_1) \mathbf{1} \\ \vdots \\ (\mathbf{Q} \circ \mathbf{F}_n) \mathbf{1} \end{bmatrix}, \quad (\mathbf{F}_\ell)_{ij} = (f_S(\mathbf{u}_i, \mathbf{u}_j))_\ell.$$

- Jacobian matrix involves Jacobian of f_S

$$\frac{\partial \mathbf{f}}{\partial \mathbf{u}} = \begin{bmatrix} \mathbf{F}_{1,\mathbf{u}_1} & \dots & \mathbf{F}_{1,\mathbf{u}_n} \\ \vdots & \ddots & \vdots \\ \mathbf{F}_{n,\mathbf{u}_1} & \dots & \mathbf{F}_{n,\mathbf{u}_n} \end{bmatrix}, \quad \partial \mathbf{F}_{i,\mathbf{u}_j} \iff \frac{\partial (f_S)_i}{\partial \mathbf{u}_{R,j}}$$

$$\mathbf{F}_{i,\mathbf{u}_j} = (\mathbf{Q} \circ \partial \mathbf{F}_{i,\mathbf{u}_j}) \pm \text{diag}(\mathbf{1}^T (\mathbf{Q} \circ \partial \mathbf{F}_{i,\mathbf{u}_j}))$$

Extensions: dissipative terms

- Define anti-symmetric entropy dissipative flux (e.g., Lax-Friedrichs)

$$\begin{aligned} \mathbf{d}_S(\mathbf{u}_L, \mathbf{u}_R) &= -\mathbf{d}_S(\mathbf{u}_R, \mathbf{u}_L) \\ (\mathbf{v}_L - \mathbf{v}_R)^T \mathbf{d}_S(\mathbf{u}_L, \mathbf{u}_R) &\geq 0. \end{aligned}$$

- Dissipation matrix \mathbf{K} (symmetric, non-negative entries)

$$\mathbf{d}(\mathbf{u}) = (\mathbf{K} \circ \mathbf{D}) \mathbf{1}, \quad \mathbf{D}_{ij} = \mathbf{d}_S(\mathbf{u}_i, \mathbf{u}_j).$$

- Jacobian of $\mathbf{d}(\mathbf{u})$ is similar to previous formulas

$$\frac{\partial \mathbf{d}}{\partial \mathbf{u}} = -(\mathbf{K} \circ \partial \mathbf{D}_R^T) + \text{diag}((\mathbf{K} \circ \partial \mathbf{D}_R^T) \mathbf{1}).$$

Extensions: modal DG methods (entropy projection)

$$\frac{\partial \mathbf{r}}{\partial \widehat{\mathbf{u}}} = \begin{matrix} \mathbf{V}_h^T \\ (N_p \times N_{\text{total}}) \end{matrix} \quad \begin{matrix} \mathbf{Q} \circ \partial \mathbf{F} \\ (N_{\text{total}} \times N_{\text{total}}) \end{matrix} \quad \begin{matrix} \frac{\partial \mathbf{u}}{\partial \mathbf{v}} \Big|_{\mathbf{V}_h \mathbf{P} \mathbf{v}_q} \\ (N_{\text{total}} \times N_{\text{total}}) \end{matrix} \quad \begin{matrix} \mathbf{V}_h \mathbf{P} \\ (N_{\text{total}} \times N_q) \end{matrix} \quad \begin{matrix} \frac{\partial \mathbf{v}}{\partial \mathbf{u}} \Big|_{\mathbf{V} \widehat{\mathbf{u}}} \\ (N_q \times N_q) \end{matrix} \quad \begin{matrix} \mathbf{V} \\ (N_q \times N_p) \end{matrix}$$

- Suppose degrees of freedom are N_p “modal” coefficients $\widehat{\mathbf{u}}$.
- Fluxes use entropy projected conservative variables \mathbf{u} ($\Pi_N \mathbf{v}(\mathbf{u}_h)$).
- Jacobian requires projection/interpolation matrices and Jacobians of transformations between conservative and entropy variables.

Talk outline

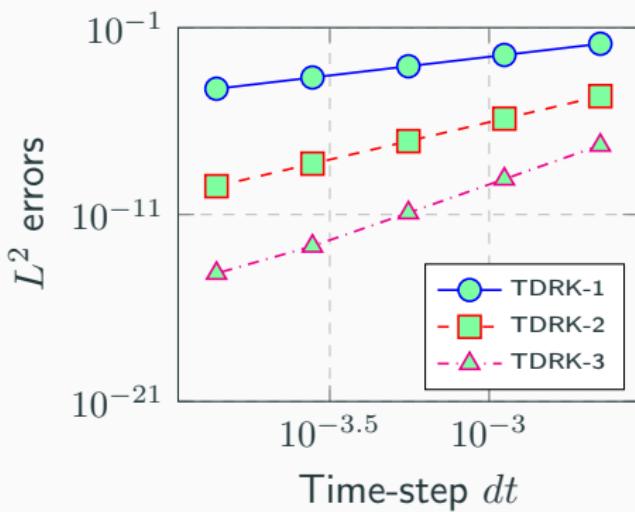
- 1 Entropy stable nodal summation-by-parts (SBP) schemes
- 2 Jacobian matrices for “flux differencing” formulations
- 3 Numerical experiments

Computational timings

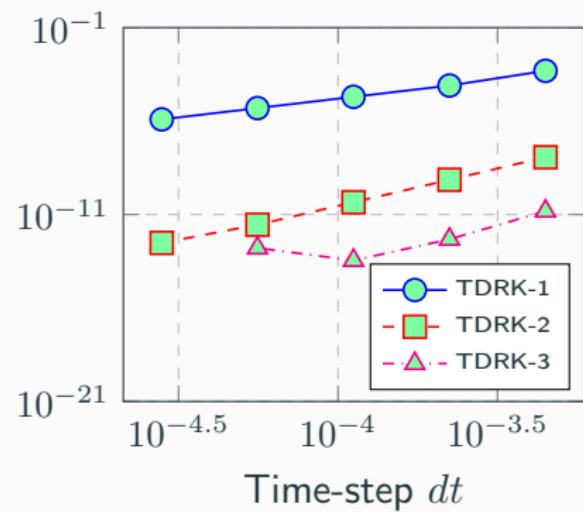
- Ratio of cost of flux evaluation to cost of AD in Julia:
 - For $1/(uL + uR)$: $7.132 \mu\text{s} / 13.593 \mu\text{s}$
 - For logmean $\{\{u\}\}^{\log} = \frac{u_L - u_R}{\log u_L - \log u_R}$: $129.254 \mu\text{s} / 161.322 \mu\text{s}$
- Jacobian timings for $f_S(u_L, u_R) = \frac{1}{6}(u_L^2 + u_L u_R + u_R^2)$ and dense differentiation matrices $\mathbf{Q} \in \mathbb{R}^{N \times N}$.

	N = 10	N = 25	N = 50
Direct automatic differentiation	5.666	60.388	373.633
FiniteDiff.jl	1.429	17.324	125.894
Jacobian formula (analytic flux deriv.)	.209	1.005	3.249
Jacobian formula (AD flux deriv.)	.210	1.030	3.259
Evaluation of $\mathbf{f}(\mathbf{u})$ (for reference)	.120	.623	2.403

Application: two-derivative time-stepping methods



(a) Burgers' equation

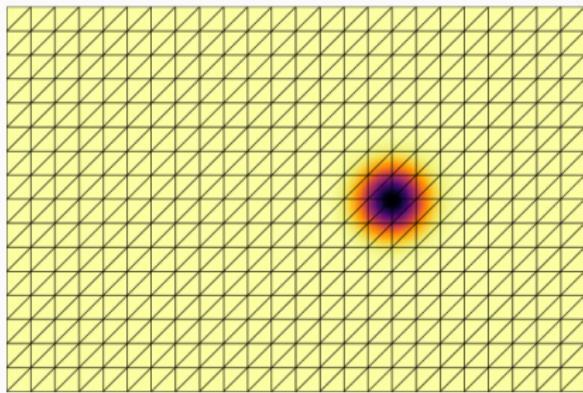


(b) Shallow water equations

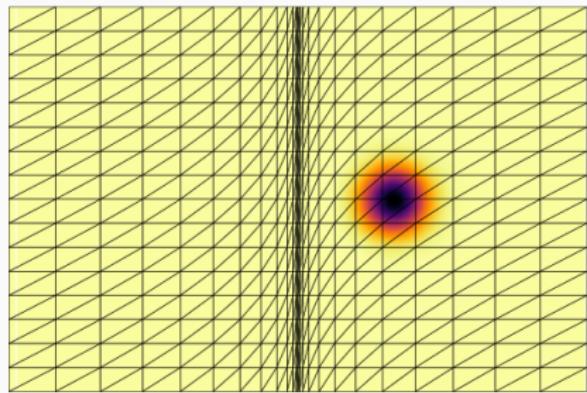
Two-derivative Runge-Kutta (TDRK) schemes: 2nd order example

$$\mathbf{u}^{k+1} = \mathbf{u}^k - \Delta t \mathbf{f}(\mathbf{u}^k) + \frac{\Delta t^2}{2} \mathbf{g}(\mathbf{u}^k), \quad \mathbf{g}(\mathbf{u}) = \frac{d^2 \mathbf{u}}{dt^2} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \frac{d\mathbf{u}}{dt} = -\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \mathbf{f}(\mathbf{u}).$$

Application: implicit midpoint method, compressible Euler



(a) Uniform mesh, L^2 error .0901



(b) Anisotropic mesh, L^2 error .0935

Figure: Solutions for a degree $N = 3$ modal DG method with $dt = .1$ on uniform and “squeezed” meshes.

Conclusions

- Simple Jacobian formulas for entropy stable flux differencing schemes.
- Concise and efficient Julia implementation `FluxDiffUtils.jl` (available on Github, will be registered soon).
- This work was supported by DMS-1719818 and DMS-CAREER-1943186.

Thank you! Questions?

