Overcoming First Order



8/10/23

Steven Roberts Sidney Fernbach Postdoctoral Fellow

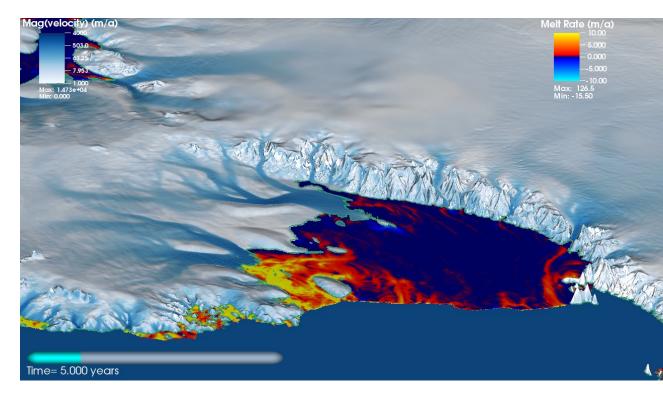


LLNL-PRES-852733 This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC



(Multi)physics Simulations Often use Low Temporal Order

- Sometimes this is fine
 - Non-smoothness
 - Low accuracy is acceptable
- Sometimes high order brings skepticism
 - Will the cost per step be too high?
 - Will the stability improve enough to offset the cost?
- Sometimes the numerical method is to blame
 - Low order splitting methods limit the overall order
 - Stiffness may cause order reduction
 - Constraints are not accurately enforced







This Talk will Focus on Runge-Kutta Methods

A Runge-Kutta method solves the ordinary differential equation (ODE)

$$y' = f(y), \quad y(t_0) = y_0$$

with the numerical procedure

$$Y_{i} = y_{n} + \Delta t \sum_{\substack{j=1\\s}}^{s} a_{i,j} f(Y_{j}), \quad i = 1, \dots, s,$$

$$\frac{C | A|}{| b^{T}|}$$

$$y_{n+1} = y_{n} + \Delta t \sum_{\substack{j=1\\j=1}}^{s} b_{j} f(Y_{j})$$





Part 1: Stiff Order Conditions and Runge-Kutta Methods for Semilinear ODEs

This work was done in collaboration with



David Shirokoff New Jersey Institute of Technology



Abhijit Biswas King Abdullah University of Science and Technology



David Ketcheson King Abdullah University of Science and Technology



Benjamin Seibold Temple University





Order Reduction Arises From Simple Problems

Prothero and Robinson¹ proposed the simple problem

$$y' = \lambda \big(y - \phi(t) \big) + \phi'(t)$$

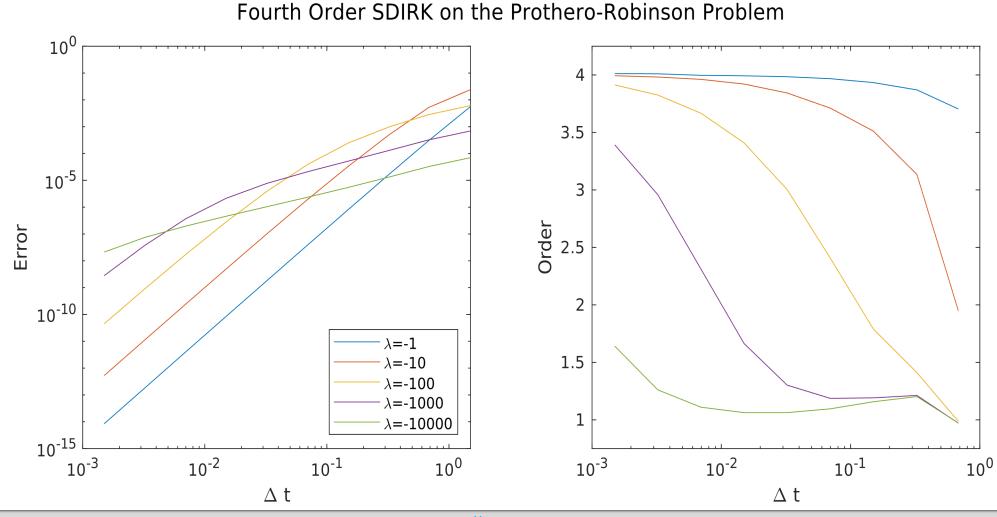
- When $|\lambda| \gg \Delta t^{-1}$, a Runge-Kutta method may converge at an order lower than the classical order.
- This phenomenon is called order reduction.
- Classical order condition theory makes unrealistic assumptions for stiff problems
 - The right-hand side has a moderate Lipschitz constant independent of Δt : $||f(y) f(z)|| \le L||y z||$
 - The time step is "sufficiently small"





^{1.} Prothero, A., and A. Robinson. "On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations." Mathematics of Computation 28.125 (1974): 145-162.

The Prothero-Robinson Problem Reveals Order Reduction



Lawrence Livermore National Laboratory





My Problem Does Not Look Like This

$$y' = \lambda \big(y - \phi(t) \big) + \phi'(t)$$

What about PDEs?







Linear PDEs also Cause Order Reduction

Much of the Prothero-Robinson convergence theory extends to linear PDEs

 $u_t = L(x, \partial)u + g(t)$

- We often see fractional convergence orders depending on $L(x, \partial)$ and the norm used¹
- Many authors have identified the additional stiff order conditions

$$0 = b^T (I - zA)^{-1} \left(Ac^{k-1} - \frac{c^k}{k} \right), \quad \forall z \in \mathbb{C}^-, k = 1, \dots, q$$

- The largest q for which this holds is the weak stage order² or pseudostage order³
 - Explicit and diagonally implicit methods can have high weak stage order

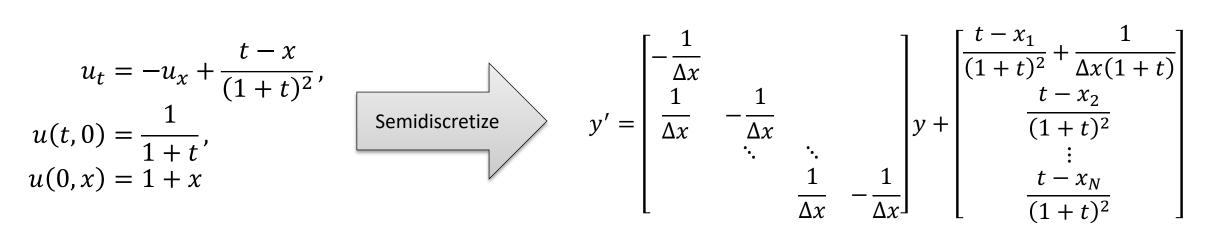
3. Skvortsov, LM. "How to avoid accuracy and order reduction in Runge-Kutta methods as applied to stiff problems." Computational Mathematics and Mathematical Physics 57 (2017): 1124-1139.



Ostermann, Alexander, and Michel Roche. "Runge-Kutta methods for partial differential equations and fractional orders of convergence." *Mathematics of Computation* 59.200 (1992): 403-420.
 Ketcheson, David I., et al. "DIRK schemes with high weak stage order." *Spectral and High Order Methods for Partial Differential Equations* (2020): 453.

What Happens on a Simple Advection PDE?

• Let's solve the following PDE¹ on $t, x \in [0,1]$:



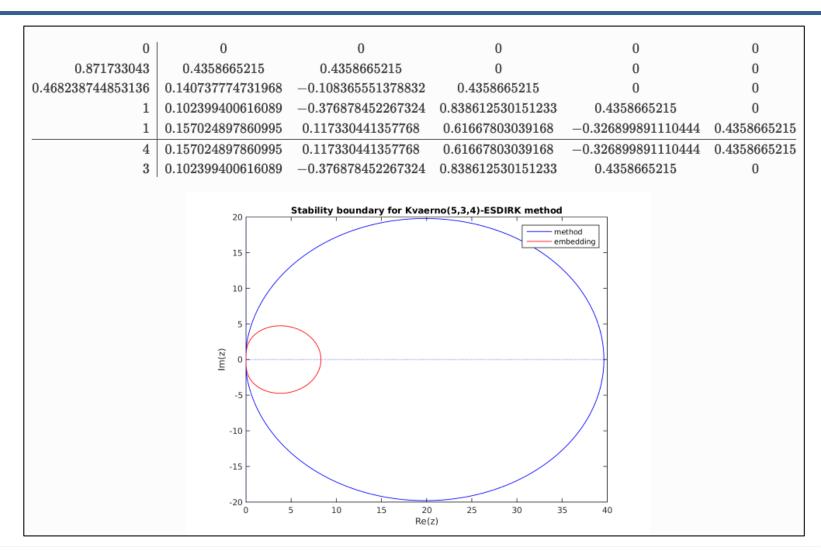
- This finite difference discretization contributes no spatial error.
- The Lipschitz constant is $L = \frac{1}{\Delta x}$, so we need $\Delta t < C\Delta x$ to be in the asymptotic regime. — This looks like a CFL condition even for implicit methods

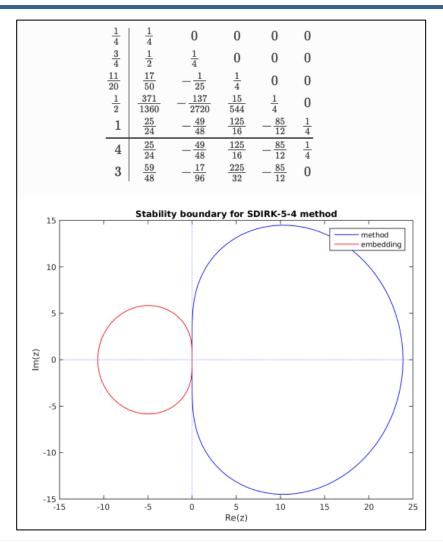




^{1.} Sanz-Serna, Jesús María, Jan G. Verwer, and W. H. Hundsdorfer. "Convergence and order reduction of Runge-Kutta schemes applied to evolutionary problems in partial differential equations." *Numerische Mathematik* 50.4 (1986): 405-418.

We Solve the Advection PDE with Two Fourth Order DIRK Methods from SUNDIALS

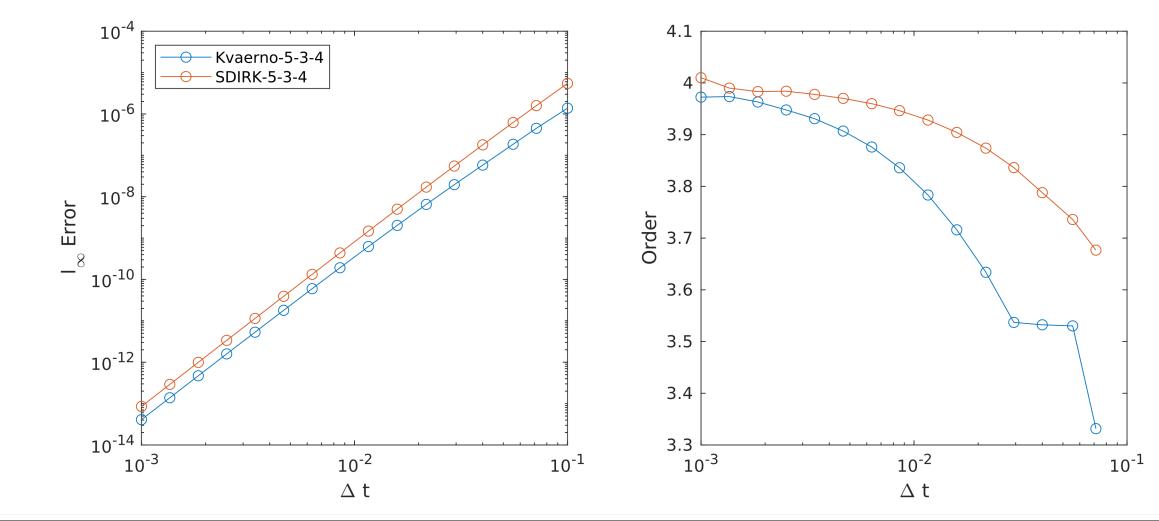








We See Asymptotic Convergence on a 16 Point Grid

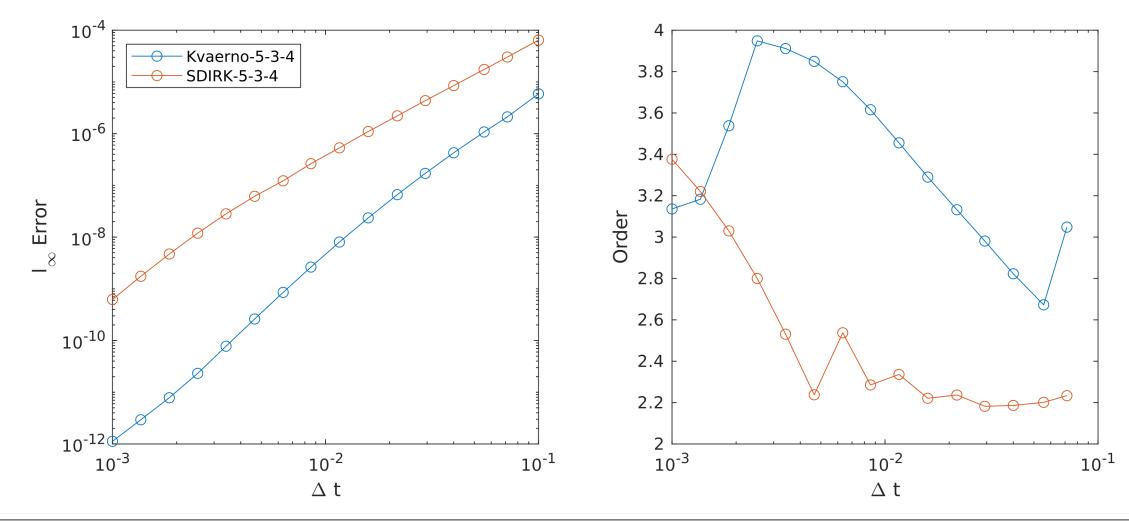


Lawrence Livermore National Laboratory LLNL-PRES-852733





We See Order Reduction on a 2048 Point Grid



Lawrence Livermore National Laboratory





My Problem Does Not Look Like This

$u_t = L(x, \partial)u + g(t)$

What about nonlinear problems?







Nonlinear Problems Require Stringent Order Conditions

- Nonlinearity often worsens order reduction
- The typical remedy is high stage order

C(q):
$$Ac^{k-1} = \frac{c^k}{k}, \quad k = 1, ..., q,$$

B(p): $b^T c^{k-1} = \frac{1}{k}, \quad k = 1, ..., p$

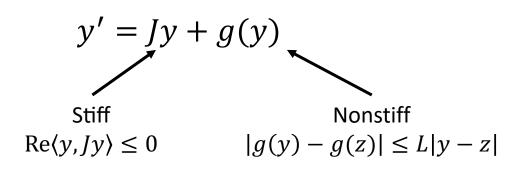
- This is very restrictive!
 - Explicit methods have max stage order of 1
 - Diagonally implicit methods have max stage order of 2
- Within the Runge-Kutta family, fully implicit schemes are seemingly the only ones that can achieve high orders outside the classical regime.



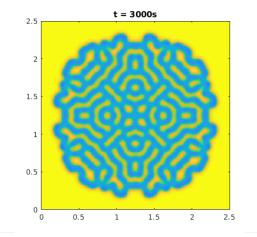


We Consider Semilinear Problems

- In nonlinear problems, stiffness often arises from linear terms
- Let's consider semilinear problems



- Examples include
 - Patten-forming diffusion reaction problems
 - Schrödinger equations
 - Air pollution transport models







The Situation for Semilinear Problems is Unclear

$$C(q): \quad Ac^{k-1} = \frac{c^k}{k}, \quad k = 1, ..., q,$$

$$B(p): \quad b^T c^{k-1} = \frac{1}{k}, \quad k = 1, ..., p$$

- Do we need the restrictive condition of high stage order for semilinear problems?
 The literature suggests yes
- Are there sharper order conditions for semilinear problems?
- Can we find methods devoid of order reduction with practical structures?
 - We will focus on diagonally implicit methods

Theorem 3.3: Let $\alpha, \beta \in \mathbb{R}$ be given. Assume the Runge-Kutta method (1.3) is A-stable, AS-stable and ASI-stable. Then we have for the class of problems (1.5) satisfying (1.6) the (optimal) B-convergence result $|\varepsilon_N| \le C \tau^p \quad (0 < \tau \le \overline{\tau})$ with order (a) p = q if B(q), C(q),

(b) p=q+ if B(q+1), C(q) and ψ is uniformly bounded on \mathbb{C}^- .

Burrage, Kevin, W. H. Hundsdorfer, and Jan G. Verwer. "A study of B-convergence of Runge-Kutta methods." *Computing* 36.1-2 (1986): 17-34.

(3.3)	$p = \left\{ egin{array}{c} q \ q+1 \end{array} ight.$	if $B(q)$ and $C(q)$ ho d, if $B(q+1)$ and $C(q)$ hold and $\psi(z)$ is uniformly bounded on \mathbb{C}^- ,	
Theore	зм 3.4.		
	Runge–Kutta metho order p given by (ods of the family \mathcal{M}_1 are convergent on the class \mathcal{F}_1 (3.3)–(3.5).	
ii) All Runge-Kutta methods of the family M₂ are convergent on the class F₂ with order p given by (3.3)-(3.5).			

Calvo, M., S. González-Pinto, and J. I. Montijano. "Runge-Kutta methods for the numerical solution of stiff semilinear systems." *BIT Numerical Mathematics* 40 (2000): 611-639.





Sharper Order Conditions Do Exists for Stiff Semilinear Problems

- We propose conditions that ensure a convergence order uniformly with respect to stiffness
- Like classical order conditions, there is 1-to-1 correspondence with rooted trees
- The semilinear order conditions contain weak stage order condition

 They coincide up to order 3
- The semilinear order conditions are sharper than stage order conditions

		Standard	
Label	Tree t	Form of t	Order Condition
1a	•	$[au^0]$	$0 = 1 - b^T \mathbb{1}$
2a		[au]	$0 = \frac{1}{2} - b^T c + z_1 b^T (I - z_1 A)^{-1} \left(\frac{c^2}{2} - Ac \right)$
3a		$[\tau^2]$	$0 = \frac{1}{6} - \frac{b^T c^2}{2} + z_1 b^T (I - z_1 A)^{-1} \left(\frac{c^3}{6} - \frac{Ac^2}{2}\right)$
3b		$[[\tau]]$	$0 = b^T (I - z_1 A)^{-1} (I - z_2 A)^{-1} \left(\frac{c^2}{2} - Ac\right)$
4a		$[au^3]$	$0 = \frac{1}{24} - \frac{b^T c^3}{6} + z_1 b^T (I - z_1 A)^{-1} \left(\frac{c^4}{24} - \frac{Ac^3}{6}\right)$
4b	Ŷ.	$[au \ [au]]$	$0 = b^T (I - z_1 A)^{-1} C (I - z_2 A)^{-1} \left(\frac{c^2}{2} - Ac\right)$
4c	Y	$[[\tau^2]]$	$0 = b^T (I - z_1 A)^{-1} (I - z_2 A)^{-1} \left(\frac{c^3}{6} - \frac{Ac^2}{2}\right)$
4d	>	$[[[\tau]]]$	$0 = b^T (I - z_1 A)^{-1} A (I - z_2 A)^{-1} (I - z_3 A)^{-1} \left(\frac{c^2}{2} - Ac\right)$





Our Error Expansion Uses Bounded Terms

• A classical expansion of the local truncation error looks like

$$y(t_1) - y_1 = \dots + \Delta t^2 \left(\frac{1}{2} - b^T c\right) \left(\underbrace{I}_{+} g'(y_0) \right) y_0 + \Delta t^3 \left(\frac{1}{6} - b^T A c\right) \left(\underbrace{I}_{+} g'(y_0) \right)^2 y_0 + \dots$$
Unbounded
Terms

Our new semilinear expansion looks like

$$y(t_{1}) - y_{1}$$

$$= \dots + \Delta t^{2} \left(\frac{1}{2} - b^{T}c + zb^{T}(I - zA)^{-1} \left(\frac{c^{2}}{2} - Ac \right) \right) y''(t_{0}) + \Delta t^{3}b^{T}(I - zA)^{-2} \left(\frac{c^{2}}{2} - Ac \right) g'(y_{0})y''(t_{0})$$

$$+ \dots$$
Bounded
Terms
where $z = \Delta t J$ (scalar here for simplicity).





Our Semilinear Analysis Extends a Lesser-Known Classical Analysis

- Butcher trees and B-series are the typical tools for analyzing the local error of a Runge-Kutta scheme
- Albrecht¹ proposed alternative order conditions that do not (necessarily) use trees
 - Order 4 conditions, for example:

$$\begin{array}{l} \bullet \ 0 = \frac{1}{24} - \frac{b^{T}c^{3}}{6} \\ \bullet \ 0 = b^{T}C\left(\frac{c^{2}}{2} - Ac\right) \\ \bullet \ 0 = b^{T}\left(\frac{c^{3}}{6} - \frac{Ac^{2}}{2}\right) \\ \bullet \ 0 = b^{T}A\left(\frac{c^{2}}{2} - Ac\right) \end{array}$$

$$\begin{array}{l} 0 = \frac{1}{24} - \frac{b^{T}c^{3}}{6} + z_{1}b^{T}(I - z_{1}A)^{-1}\left(\frac{c^{4}}{24} - \frac{Ac^{3}}{6}\right) \\ 0 = b^{T}(I - z_{1}A)^{-1}C(I - z_{2}A)^{-1}\left(\frac{c^{2}}{2} - Ac\right) \\ 0 = b^{T}(I - z_{1}A)^{-1}(I - z_{2}A)^{-1}\left(\frac{c^{3}}{6} - \frac{Ac^{2}}{2}\right) \\ 0 = b^{T}(I - z_{1}A)^{-1}A(I - z_{2}A)^{-1}(I - z_{3}A)^{-1}\left(\frac{c^{2}}{2} - Ac\right) \end{array}$$

• We closely follow Albrecht's derivation, but the stiff, linear term introduces extra factors of $(I - hA \otimes J)^{-1}$.





^{1.} Albrecht, Peter. "The Runge–Kutta theory in a nutshell." SIAM Journal on Numerical Analysis 33.5 (1996): 1712-1735.

Practical Methods with Semilinear Order 3 Exist

- Desired properties
 - Order 3
 - Singly diagonally implicit
 - L-stable
- Typically, this requires at least 3 stages
- With order 3 semilinear conditions, this requires at least 5 stages

<u>13</u> 58	<u>13</u> 58	Θ	0	Θ	Θ	0
<u>26</u> 29	<u>39</u> 58	<u>13</u> 58	0	Θ	Θ	0
Θ	$-\frac{13}{58}$	Θ	<u>13</u> 58	Θ	Θ	Θ
<u>13</u> 29	<u>65</u> 174	$-\frac{13}{348}$	$-\frac{13}{116}$	<u>13</u> 58	Θ	0
<u>12 971</u>	2 015 824 758 301 938 982 625	554 819 849 934 875	68 790 302 177 688 571 375	7 705 505 568 680 430 000	<u>13</u>	0
17 611	11 720 872 553 456 507 801 646	11 076 945 065 425 668	269 445 346 056 471 443 716	56 998 053 973 484 343 863	58	
1	3 455 277 656	1061001132073	780 513 524 467	342 906 676 217	77 214 825 271 310 213 828 561	<u>13</u>
-	28 312 464 375	3 749 092 092 720	5 751 892 408 080	1 125 548 760 960	155 527 924 398 245 799 120 000	58
	3 455 277 656	1061001132073	780 513 524 467	342 906 676 217	77 214 825 271 310 213 828 561	13
	28 312 464 375	3 749 092 092 720	5 751 892 408 080	1 125 548 760 960	155 527 924 398 245 799 120 000	58
	83 396 117 862 679 251 596 686	51 873 391 680 781 295 917 121	91 834 777 272 491 463 252 761	5 676 271 777 638 433 424 524	<u>11</u>	2
	543 808 069 678 473 491 279 817	197 748 388 973 990 360 465 388	725 077 426 237 964 655 039 756	20 141 039 617 721 240 417 771	23	9

CASC



Higher Order Methods are Challenging to Derive

The first nonlinear order condition appears at order 4

$$0 = b^{T} (I - z_{1}A)^{-1} C (I - z_{2}A)^{-1} \left(\frac{c^{2}}{2} - Ac\right)$$

Neumann
Expansion

$$0 = b^{T} A^{i} C A^{j} \left(\frac{c^{2}}{2} - Ac\right), \quad i, j = 0, ..., s - 1$$

- The number or order conditions increases with the order and the number of stages
- Explicit methods of semilinear order 4 exist
- Singly diagonally implicit method of classical order 4 and semilinear order 3 exist
 - This suffices for a global order of 4
 - Methods with semilinear order 4 almost certainly exist





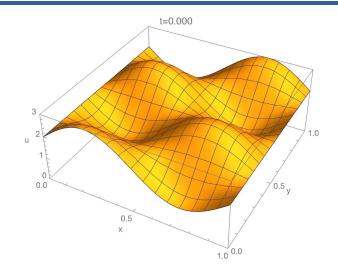


Allen-Cahn is a Semilinear PDE Modeling Phase Separation

We consider a 2D Allen-Cahn reaction-diffusion PDE

$$\frac{\partial u}{\partial t} = \alpha \nabla^2 u + \beta (u - u^3) + s(t, x, y)$$

We test methods of order 3 and 4 to validate the semilinear order conditions

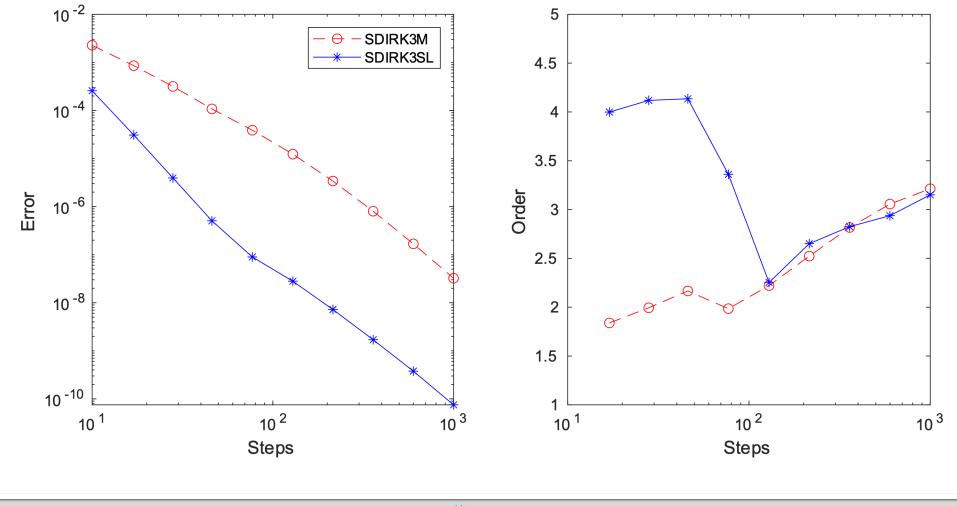


Method	Source	Stages	Classical Order	Semilinear Order
SDIRK3SL	This work	6	3	<mark>3</mark>
SDIRK3M	Kennedy, Christopher A., and Mark H. Carpenter. Diagonally implicit Runge- Kutta methods for ordinary differential equations. A review. 2016.	4	3	1
DIRK-(7,4,4)	Biswas, Abhijit, et al. "Design of DIRK schemes with high weak stage order." <i>Communications in Applied Mathematics and Computational Science</i> 18.1 (2023): 1-28.	7	4	<mark>3</mark>
SDIRK4M	Kennedy, Christopher A., and Mark H. Carpenter. Diagonally implicit Runge- Kutta methods for ordinary differential equations. A review. 2016.	5	4	1





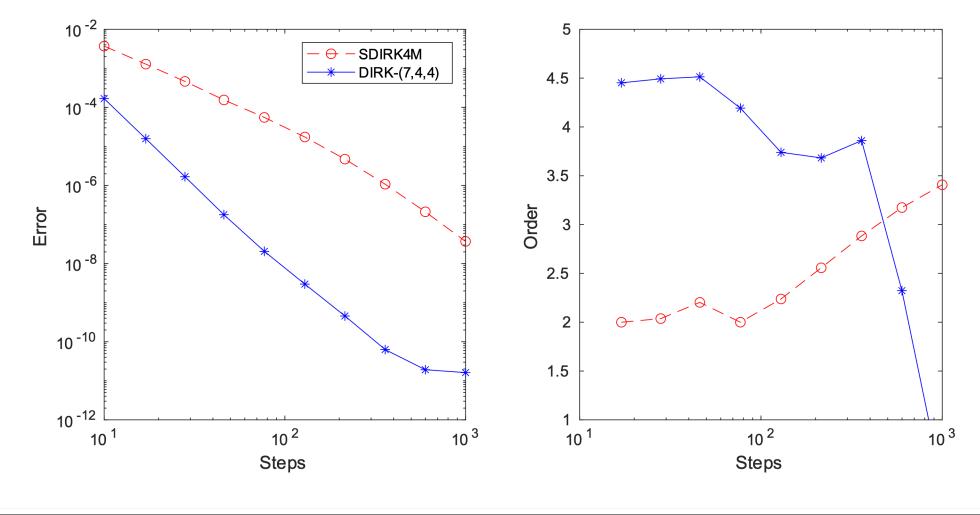
The New Method SDIRK3SL Avoid Order Reduction







The Existing DIRK-(7,4,4) Method Avoids Order Reduction







Conclusions

- Classical order conditions rely on assumptions that fail to hold for stiff problems
- The consequence is a reduction in order and accuracy for many integrators
- High stage order is not necessary to avoid order reduction on stiff, semilinear ODEs
- Weak stage order conditions for stiff, linear problems suffice up to order 4
- Order reduction and techniques to eliminate it are not limited to implicit Runge-Kutta methods
 - Implicit-explicit
 - Multirate
 - Explicit methods





Part 2: Time-Stepping for the BISICLES Ice Sheet Model

This work was done in collaboration with



Daniel Martin Lawrence Berkeley National Laboratory



David Gardner Lawrence Livermore National Laboratory



Hans Johansen Lawrence Berkeley National Laboratory



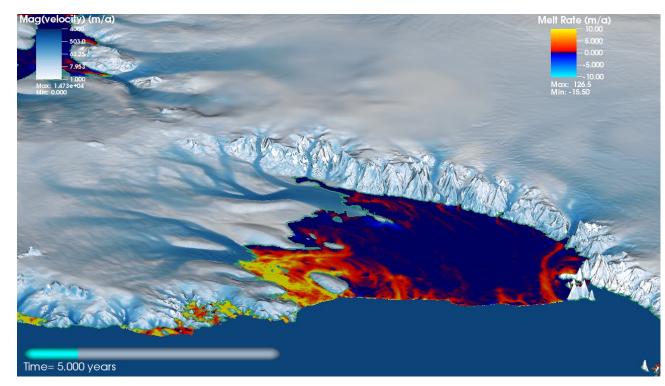
Carol Woodward Lawrence Livermore National Laboratory





BISICLES Models Ice Sheet Dynamics

- Accurate modeling of ice-sheets is critical to understanding and predicting
 - Future see level rise
 - Potential regional collapses in the West Antarctic ice sheet
- BISICLES is a simulation tool developed at LBNL, LANL, and the University of Bristol¹
- Long-term time evolution of these models requires accurate, conservative, and stable numerical methods



1. Cornford, Stephen L., et al. "Adaptive mesh, finite volume modeling of marine ice sheets." *JCP* 232.1 (2013): 529-549.



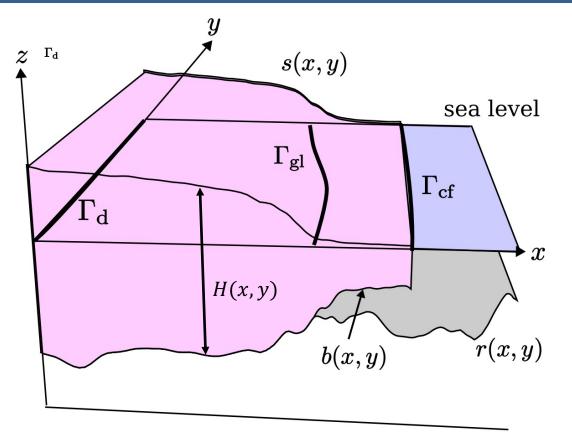


The Ice Model Combines Hyperbolic and Elliptic Partial Differential Equations

- In the simplest case, the two primary variables are
 - Ice thickness H(t, x, y)
 - Ice velocity v(t, x, y)
- An asymptotically-derived approximation to Stokes Flow is used¹

$$\frac{\partial H}{\partial t} = \frac{\partial}{\partial x}(v_x H) + \frac{\partial}{\partial y}(v_y H)$$
$$\beta^2 v - \nabla \cdot (H\mu(v) \nabla v) = -\rho_i g H \nabla \cdot s$$

- The Chombo library is used for the spatial discretization with adaptive mesh refinement (AMR)
 - Second order finite volume method



From Cornford, Stephen L., et al. "Adaptive mesh, finite volume modeling of marine ice sheets." *JCP* 232.1 (2013): 529-549.

1. Schoof, Christian, and Richard CA Hindmarsh. "Thin-film flows with wall slip: an asymptotic analysis of higher order glacier flow models." *QJMAM* 63.1 (2010): 73-114.

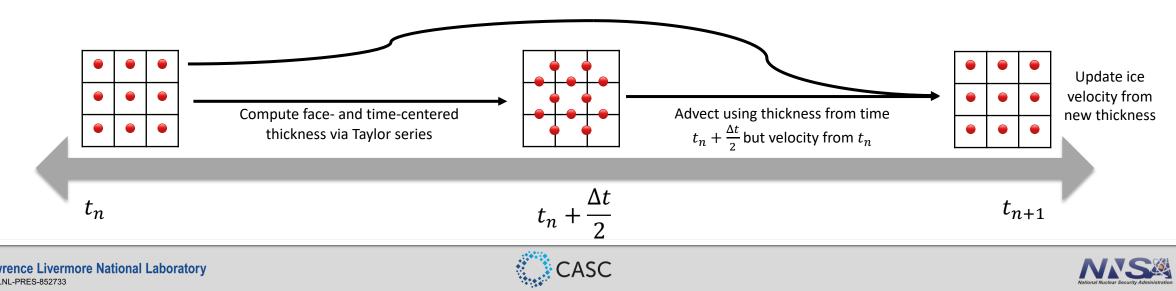




BISICLES was Limited by the Time Discretization

- BISICLES uses an unsplit Godunov piecewise parabolic method
 - First order accurate in time
 - Explicit
 - Not method of lines
 - Limited maximum stable time step
 - No error estimation
 - Time step chosen by CFL condition

- Project Goals
 - Introduce high order time-stepping methods for improved accuracy and stability
 - Introduce adaptive methods
 - Determine which class of integrators is bestsuited to the problem



We can Solve an Ordinary or Differential-Algebraic Equation

The time evolution problem is an index-1 differential-algebraic equation (DAE)

$$\frac{dH}{dt} = f(H, v)$$
$$0 = g(H, v)$$

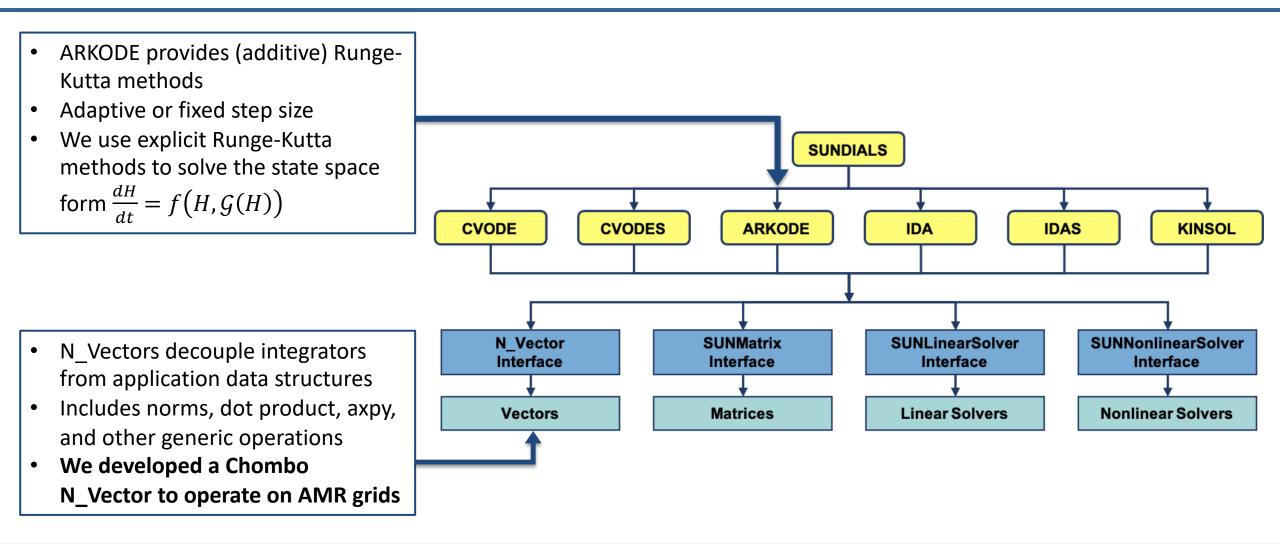
- Over 90% of the runtime is spent solving the nonlinear system 0 = g(H, v)!
- Or we can view this as an ordinary differential equation (ODE) where v = G(H) is a derived quantity computed via a nonlinear solve. This is the "state space form"

$$\frac{dH}{dt} = f(H, \mathcal{G}(H))$$





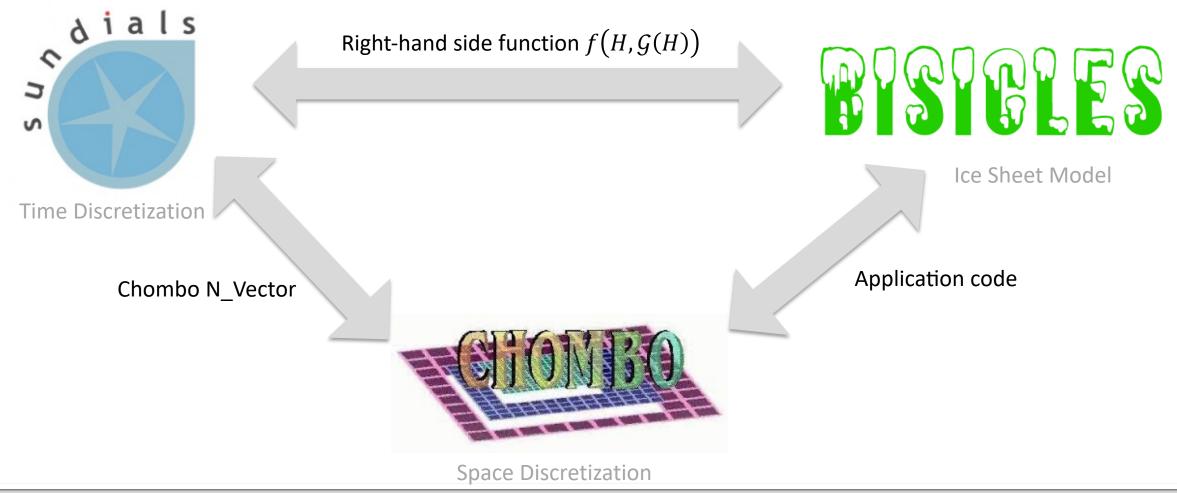
SUNDIALS Provides Efficient ODE, DAE, and Nonlinear Solvers







Our New Chombo N_Vector Enables Package Interoperability







Second Order is Feasible at the Cost of Order One

- $\frac{dH}{dt} = f(H, v)$ 0 = g(H, v)
- BISICLES' first order integrator does **1** expensive algebraic solve for ice velocity each time step
- A second order, explicit Runge-Kutta applied to the state space form requires at least 2 algebraic solves per time step
- We proposed to use a second order "half-explicit Heun's method" with 1 algebraic solve per step

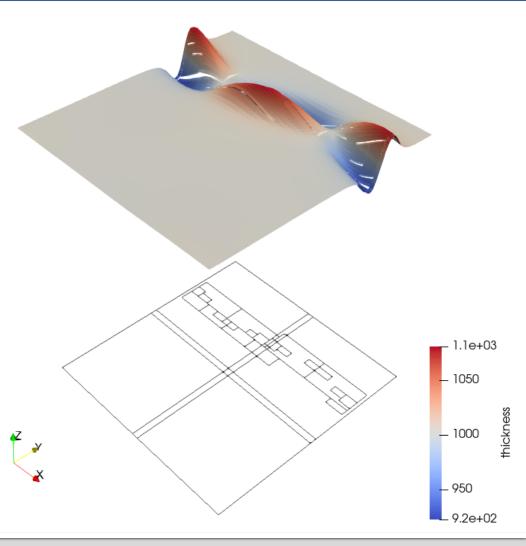
It is not a traditional Runge-Kutta method but a generalized additive Runge-Kutta.





Twisty Steam is a Benchmark Test Problem

- Ice streams are fast-flowing regions within a sheet
- Ice streams account for about 90% of ice mass lost from the Antarctic ice sheet¹
- We compare the temporal accuracy of
 - The original unsplit Godunov piecewise parabolic method in BISICLES
 - Explicit Runge-Kutta methods from ARKODE of order 1-4
 - The half-explicit Heun's method from the previous slide

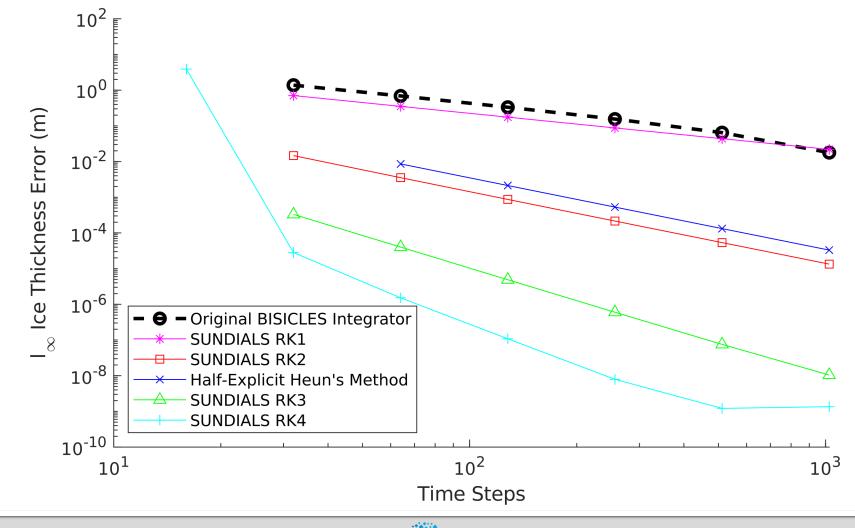


1. https://www.antarcticglaciers.org/glacier-processes/glacier-types/ice-streams/

CASC



The Integrators Converge

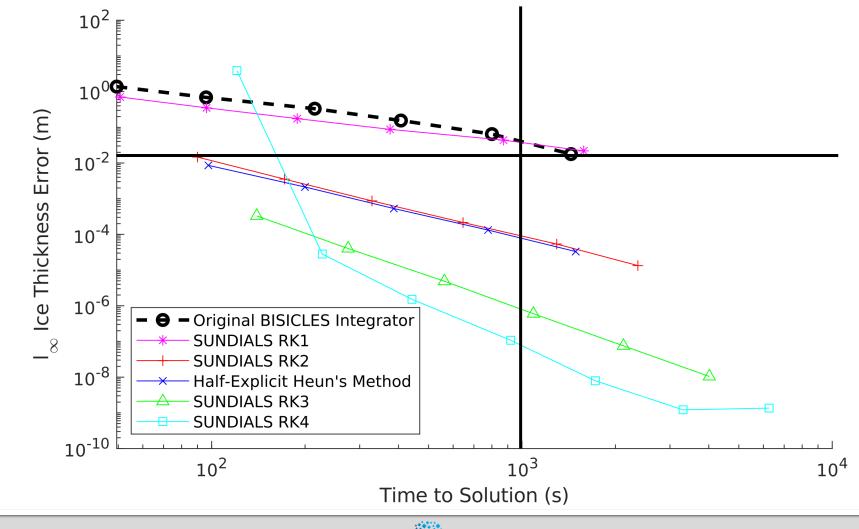


Lawrence Livermore National Laboratory





The New Integrators Are Significantly More Efficient







BISICLES is Often Stability-Limited

$$\frac{\partial H}{\partial t} = \frac{\partial}{\partial x}(v_x H) + \frac{\partial}{\partial y}(v_y H)$$
$$\beta^2 v - \nabla \cdot (H\mu(v) \nabla v) = -\rho_i g H \nabla \cdot s$$

- Despite the hyperbolic PDE for ice thickness, the problem is sometimes diffusive
- Spatial error often dominates temporal error, even when using the native, first order method
- In this regime, we achieve the best efficiency by taking Δt near the CFL limit
- The following metric is key

 $\frac{\max \text{ stable } \Delta t}{\operatorname{cost per step}}$



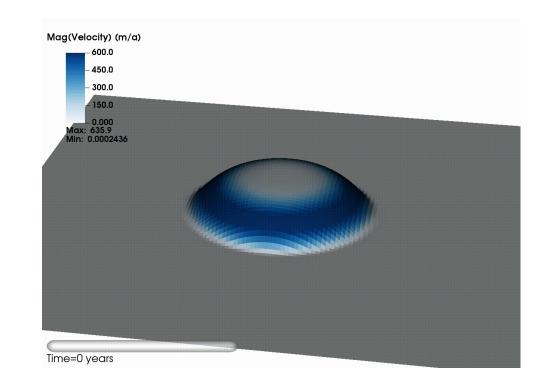


BISICLES is Often Stability-Limited

- Despite the hyperbolic PDE for ice thickness, the problem is sometimes diffusive
- Spatial error often dominates temporal error, even when using the native, first order method
- In this regime, we achieve the best efficiency by taking Δt near the CFL limit
- The following metric is key

 $\frac{\max \text{ stable } \Delta t}{\operatorname{cost per step}}$

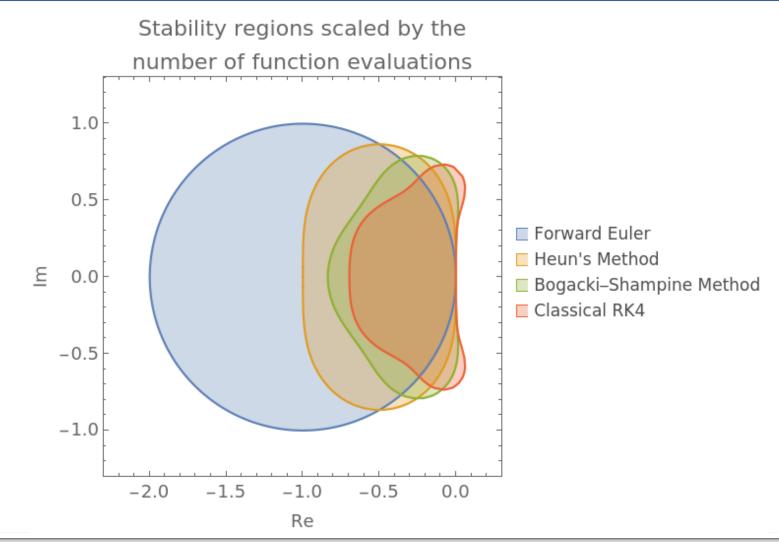
 $\frac{\partial H}{\partial t} = \frac{\partial}{\partial x} (v_x H) + \frac{\partial}{\partial y} (v_y H)$ $\beta^2 v - \nabla \cdot (H\mu(v) \nabla v) = -\rho_i g H \nabla \cdot s$







High Order is Not Always Advantageous for Linear Stability





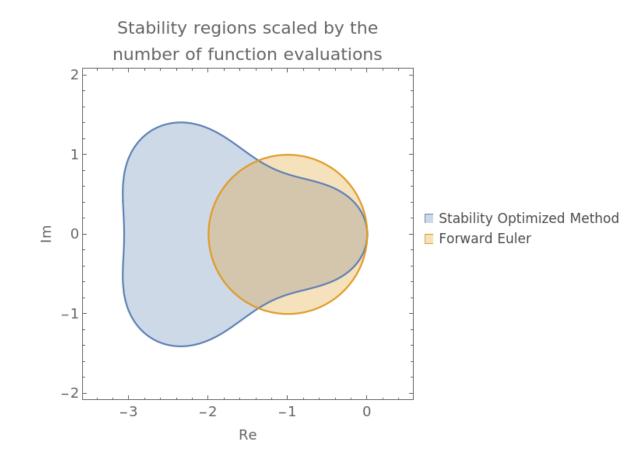


We Can Optimize The Stability with Additional Stages

 We derived a first order method in 3 stages with a large stability region

0	0	0	0
1 3 2 3	<u>1</u> 3	0	Θ
<u>2</u> 3	<u>11</u> 21	$\frac{1}{7}$	Θ
	<u>53</u> 80	<u>3</u> 40	<u>21</u> 80

- For the twisty stream problem, we can take a time step roughly 5x bigger
- The minimum time to a stable solution is reduced by about 35% for the twisty stream problem







Conclusions

- New Runge-Kutta integrators from SUNDIALS facilitate faster and more accurate modeling of ice sheets
- Embedded error estimation offers a simpler and effective alternative to CFL based time step selection
- Chombo N_Vector is now available in Chombo 3.2 patch 8
- Future and ongoing work
 - Testing multirate methods
 - Exploring other stabilized methods
 - Parallel-in-time leveraging SUNDIALS' wrappers for XBraid
 - Exploring more-complex (realistic) ice sheet configurations (grounding-line retreat, realistic Greenland and Antarctic geometries, etc).





Acknowledgements



This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Scientific Discovery through Advanced Computing (SciDAC) Program through the FASTMath Institute



computing.llnl.gov/sundials

This work was supported by the Fernbach Fellowship through the LLNL-LDRD Program under Project No. 23-ERD-048







Lawrence Livermore National Laboratory

Center for Applied

Scientific Computing

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.