A Multirate Approach to Accelerating Time-Steppers with Surrogate Models

SciCADE

CASC
Center for Applied Scientific Computing

July 27, 2022

Steven Roberts, Andrey A. Popov, Arash Sarshar, and Adrian Sandu
Sidney Fernbach Postdoctoral Fellow

LLNL-PRES-837872
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

Lawrence Livermore National Laboratory

Goal: solve large-scale initial value problems

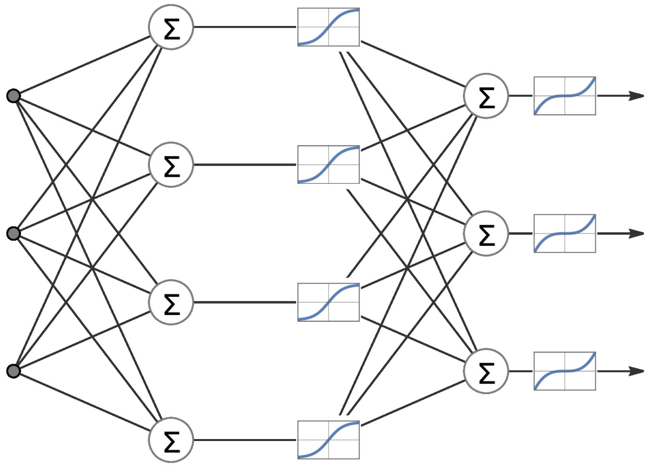
- Consider the initial value problem

$$y' = f(y), \quad y(t_0) = y_0 \in \mathbb{C}^N, \quad t \in [t_0, t_f].$$

- We will focus on explicit methods for nonstiff problems.
- In scientific applications, the dimension N can be intractably large and evaluations of f prohibitively expensive.
- How can we reduce the number of evaluations of f without sacrificing
  - Accuracy
  - Stability
  - Convergence

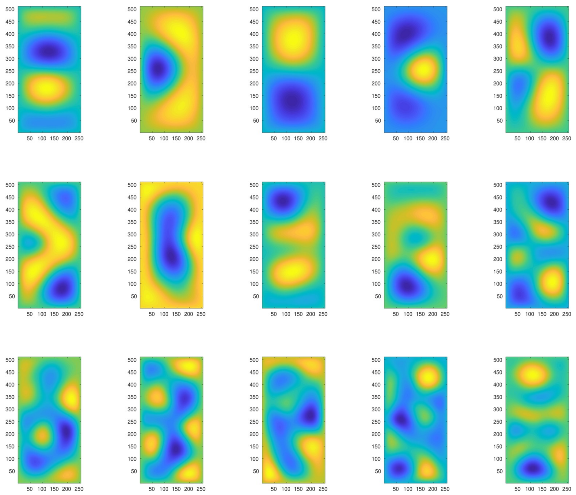
What about surrogate models?

- For many problems, it is possible to produce a cheap but approximate surrogate model.



Machine Learning

Coarser Mesh



Reduced-Order Model

- For complex problems, surrogate models cannot outright replace the full model f.

How can we combine full and surrogate models?

- For convergence, we cannot escape evaluating the full model.
- An ideal hybrid approach would use the surrogate model to substantially reduce evaluations of the full model.
- Surrogate models have been successfully incorporated into optimization algorithms.
- There are some related ideas in the context of time integration
  - Rosenbrock-W methods
  - Coupling a reduced order model and multirate method[1]
  - Mixed Precision Runge-Kutta methods[2]
  - Defect correction
  - Heterogeneous multiscale method[3]

1. Hachtel, Christoph, et al. "Multirate DAE/ODE-simulation and model order reduction for coupled field-circuit systems." Scientific Computing in Electrical Engineering. Springer, Cham, 2018. 91-100.
2. Grant, Zachary J. "Perturbed Runge–Kutta Methods for Mixed Precision Applications." *Journal of Scientific Computing* 92.1 (2022): 1-20.
3. Abdulle, Assyr, et al. "The heterogeneous multiscale method." Acta Numerica 21 (2012): 1-87.

Defining the surrogate model

- Recall the full model we want to integrate is

$$y' = f(y), \quad y(t) \in \mathbb{C}^N.$$

- The surrogate model is also posed as an ODE:

$$y'_{sur} = f_{sur}(y_{sur}), \quad y_{sur}(t) \in \mathbb{C}^S.$$

- The surrogate model may evolve in a lower-dimensional space: $S < N$.
- Transformations between the full and surrogate spaces are realized by $V, W \in \mathbb{C}^{N \times S}$:

$$y_{sur} = W^* y, \quad y \approx V y_{sur}, \quad W^* V = I_{S \times S}.$$

Surrogate acceleration with multirate methods

- The original, full ODE can be rewritten in the equivalent form

$$y' = V f_{sur}(W^* y) + f(y) - V f_{sur}(W^* y) \in \mathbb{C}^N.$$

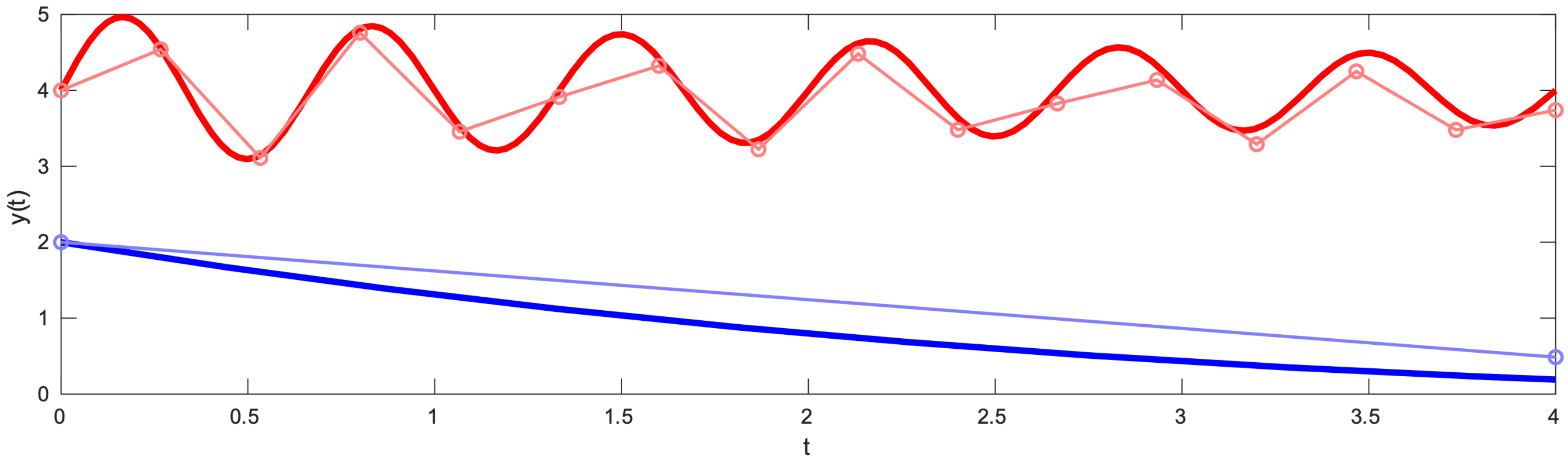
- Idea: apply a multirate method to this ODE.
  - The "fast" partition is the surrogate model and is treated with a small timestep.
  - The "slow" partition is the surrogate error and is treated with a large timestep.

- The surrogate model is evaluated often to guide the solution trajectory while the expensive full model is evaluated infrequently to correct for surrogate errors.

- Accuracy, stability, and convergence properties are based on the underlying multirate method.

Which multirate methods should we use?

- With 6 decades of development, there are many options!
- Any method suitable for additively-partitioned systems should suffice.
- Multirate infinitesimal (MRI) methods have gain traction in recent years.
  - Fast dynamics are evolved by solving ODEs with any consistent integrator.
  - Very flexible



Multirate infinitesimal Euler example

- Our multirate ODE is

$$y' = f^{\{f\}}(y) + f^{\{s\}}(y).$$

- Consider the simple multirate infinitesimal method

$$\begin{aligned} v(0) &= y_n, \\ v'(\theta) &= f^{\{f\}}\big(v(\theta)\big) + f^{\{s\}}(y_n), \\ y_{n+1} &= v(H). \end{aligned}$$

- There is one evaluation of $f^{\{s\}}$ per step.
- $f^{\{f\}}$ is evaluated as many times as it takes to integrate v to $\theta = H$.

Multirate infinitesimal Euler example

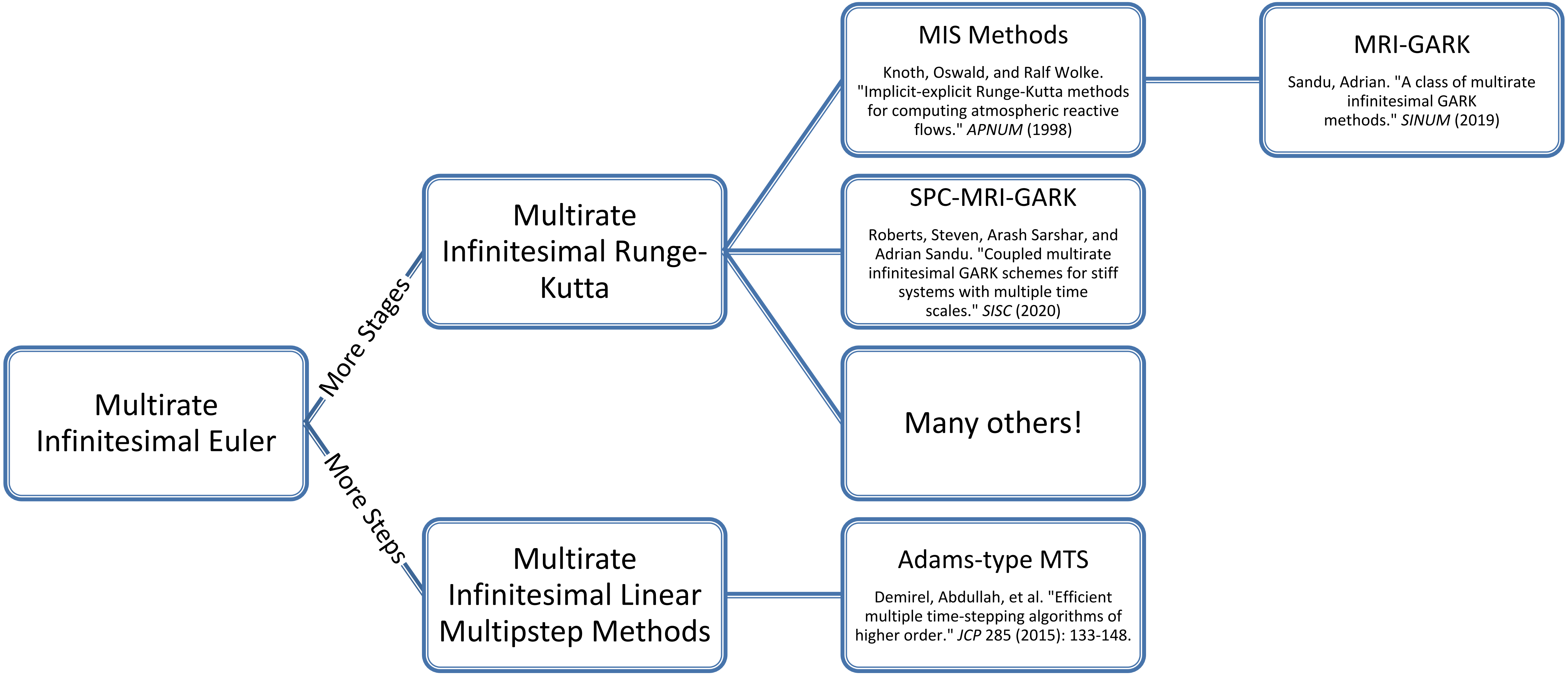
$y' = V f_{sur}(W^* y) + f(y) - V f_{sur}(W^* y)$

- When we apply the multirate Euler method to our ODE, we arrive at

$$
\begin{aligned}
z(0) &= W^* y_n, \\
z'(\theta) &= f_{sur}\big(z(\theta)\big) + W^* f(y_n) - f_{sur}(W^* y_n), \\
y_{n+1} &= V\, z(H) + (I_{N\times N} - V W^*)\big(y_n + H\, f(y_n)\big).
\end{aligned}
$$

- $z(\theta) \in \mathbb{C}^S$ is integrated in the range of V.
- An Euler step is taken in the nullspace of W^*.
- There is one evaluation of the full model per step and many for the surrogate model.

General frameworks for multirate infinitesimal methods



SM-MRI-GARK

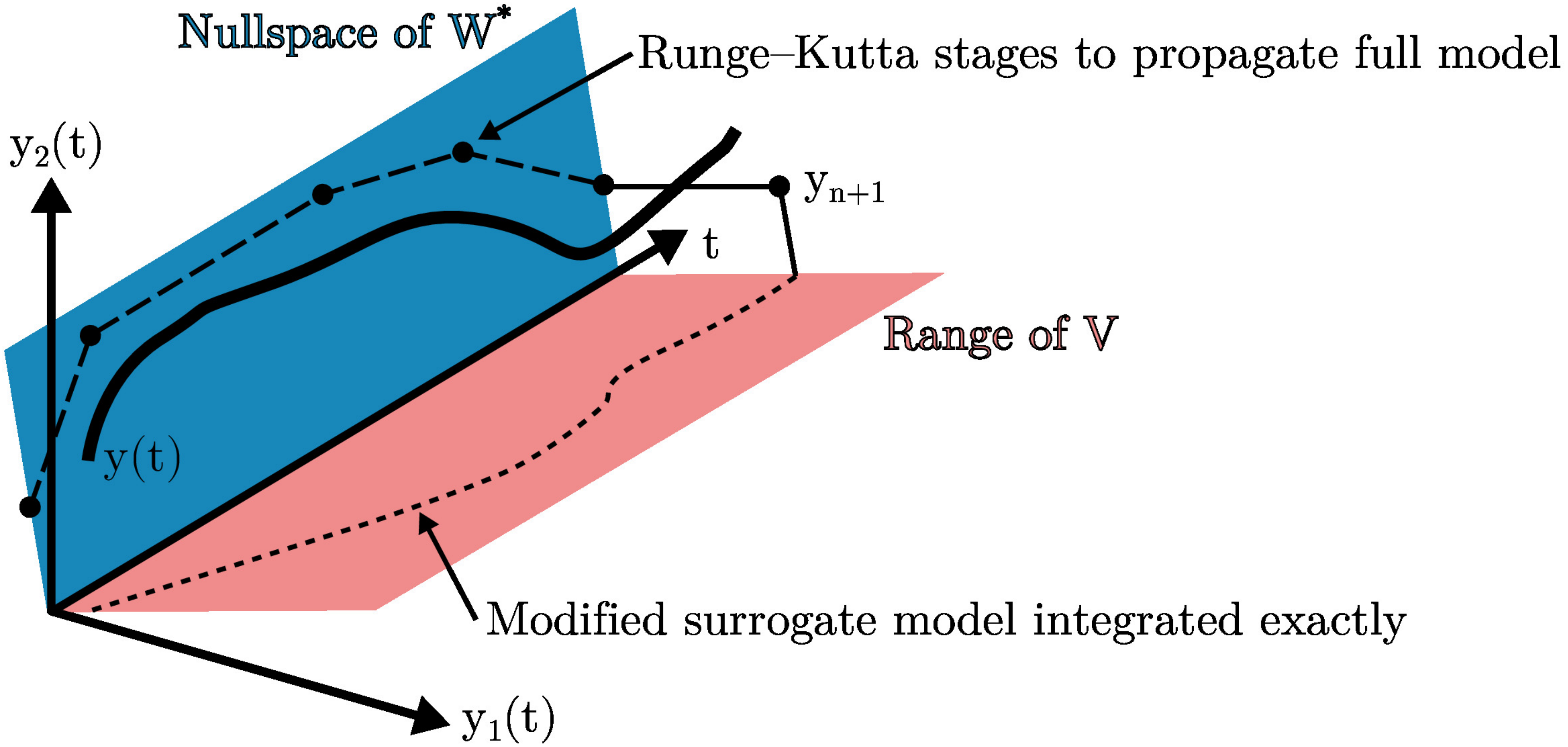
$$y' = V f_{sur}(W^* y) + f(y) - V f_{sur}(W^* y)$$

- Let's replace multirate Euler with MRI-GARK to achieve higher orders.
- A surrogate model MRI-GARK (SM-MRI-GARK)[1] method is given by

$$\begin{aligned}
Y_1 &= y_n, \\
z_i(0) &= W^* Y_i \in \mathbb{C}^S, \\
z_i'(\theta) &= \Delta c_i^{\{s\}} f_{sur}\left(z_i(\theta)\right) + \sum_{j=1}^{i+1} \gamma_{i,j}\left(\frac{\theta}{H}\right)\left(W^* f\left(Y_j\right) - f_{sur}\left(W^* Y_j\right)\right), \\
Y_{i+1} &= V\, z_i(H) + (I_{N\times N} - V W^*)\left(Y_i + H \sum_{j=1}^{i+1} \bar{\gamma}_{i,j} f\left(Y_j\right)\right), \quad i = 1, \dots, s^{\{s\}}, \\
y_{n+1} &= Y_{s^{\{s\}}+1}.
\end{aligned}$$

1. Roberts, Steven, et al. "A Fast Time-Stepping Strategy for Dynamical Systems Equipped with a Surrogate Model." *SIAM Journal on Scientific Computing* 44.3 (2022): A1405-A1427.

Illustration of the time-stepping approach



SM-SPC-MRI-GARK

$$y' = V f_{sur}(W^* y) + f(y) - V f_{sur}(W^* y)$$

- If instead we base our method on SPC-MRI-GARK we have the class of surrogate model SPC-MRI-GARK (SM-SPC-MRI-GARK)[1]:

$$Y_i = y_n + H \sum_{j=1}^{s^{\{s\}}} a_{i,j}^{\{s\}} f\left(Y_j\right), \quad i = 1, \dots, s^{\{s\}},$$

$$z(0) = W^* y_n \in \mathbb{C}^S,$$

$$z'(\theta) = f_{sur}\left(z(\theta)\right) + \sum_{j=1}^{s^{\{s\}}} \gamma_j\left(\frac{\theta}{H}\right)\left(W^* f\left(Y_j\right) - f_{sur}\left(W^* Y_j\right)\right),$$

$$y_{n+1} = V\, z(H) + (I_{N\times N} - V W^*)\left(y_n + H \sum_{j=1}^{s^{\{s\}}} b_j f\left(Y_j\right)\right)$$

1. Roberts, Steven, et al. "A Fast Time-Stepping Strategy for Dynamical Systems Equipped with a Surrogate Model." *SIAM Journal on Scientific Computing* 44.3 (2022): A1405-A1427.

SM-MRI-AB

$$y' = Vf_{sur}(W^*y) + f(y) - Vf_{sur}(W^*y)$$

- Finally, an Adams-type MTS method yields a surrogate model MRI Adams-Bahsforth (SM-MRI-AB):

$$z(0) = W^* y_n,$$

$$z'(\theta) = f_{sur}\big(z(\theta)\big) + \sum_{j=0}^{k-1} \gamma\left(\frac{\theta}{H}\right) \nabla^j \big(W^* f(y_n) - f_{sur}(y_n)\big),$$

$$y_{n+1} = v(H) + (I_{N\times N} - VW^*) \sum_{j=1}^{k-1} \bar{\gamma} \nabla^j f(y_n).$$

Numerical experiment: Lorenz ‘96

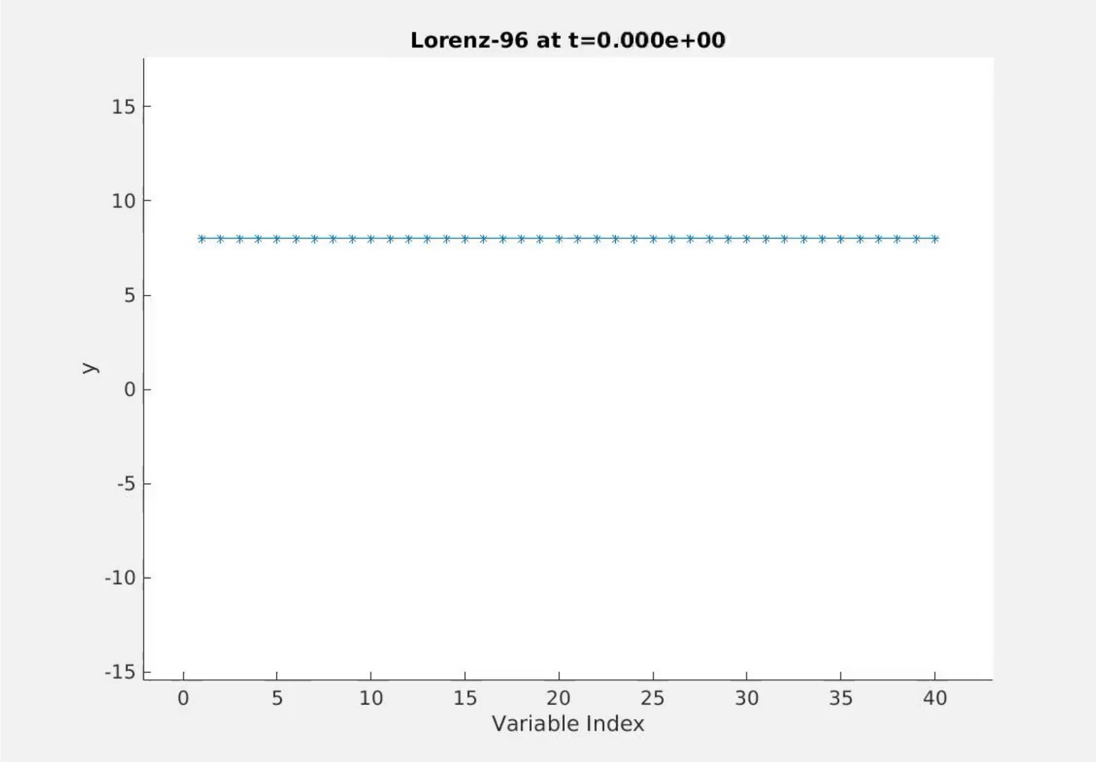
- The Lorenz ‘96 is a 40 variable ODE

$$\frac{dX_k}{dt} = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + F.$$

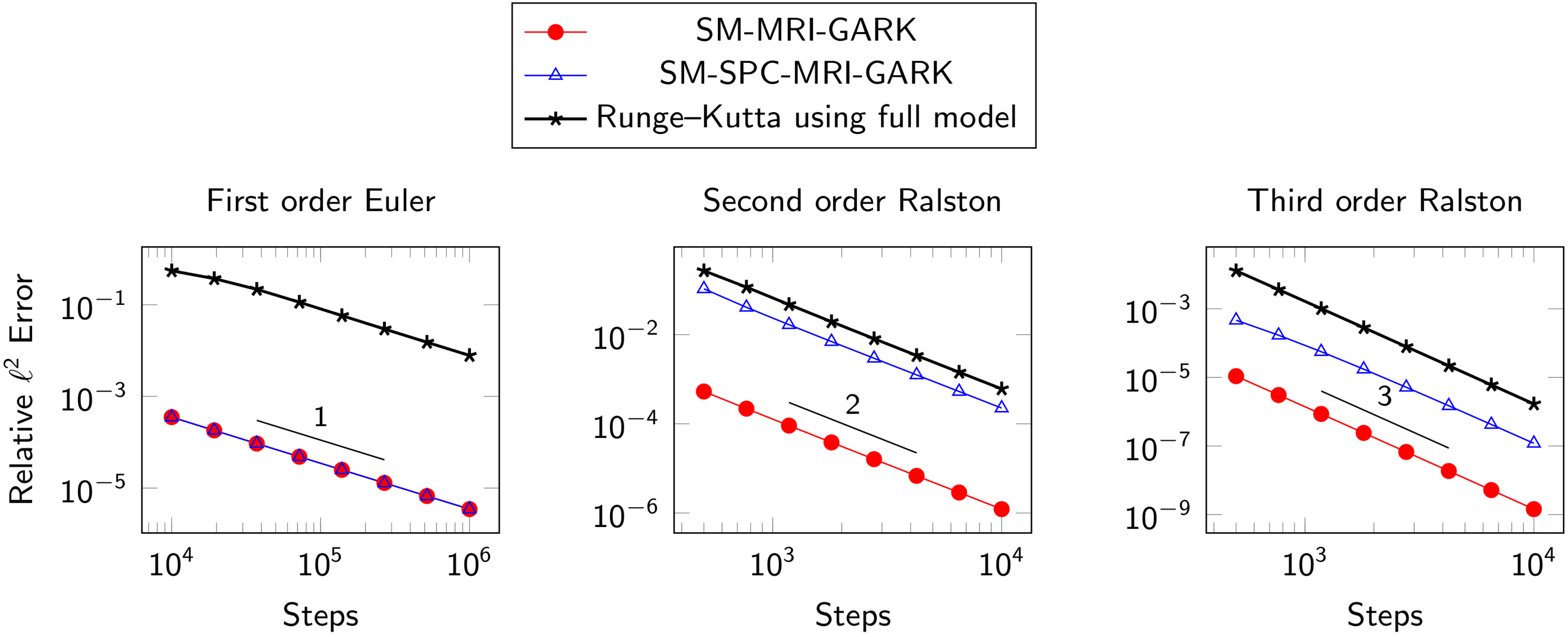
- In an offline phase, 5000 snapshots of the trajectory and its derivative were generated over the timespan $[2, 10]$.

- A 3-layer neural network was trained on the data to approximate the RHS function f.

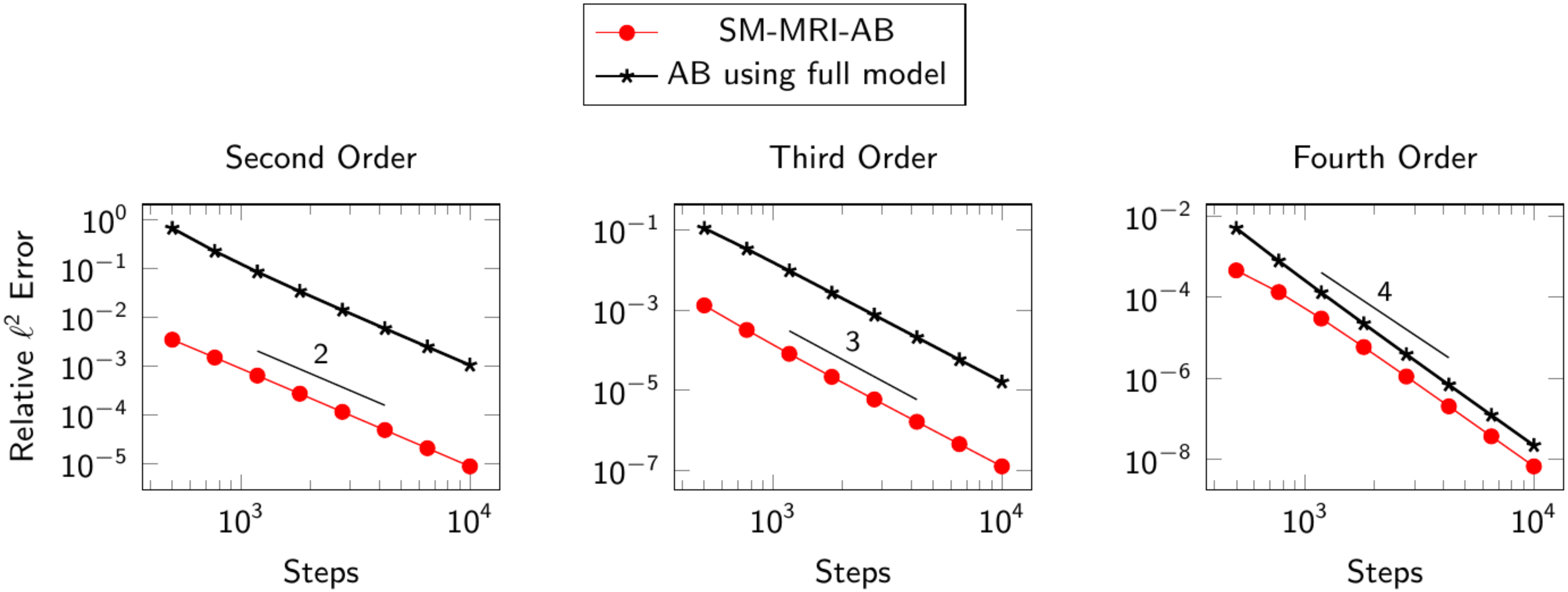
- The neural network acts as f_{sur}, and $V = W = I_{40\times40}$.



Numerical experiment: Lorenz ‘96



Numerical experiment: Lorenz '96



Numerical experiment: DG advection

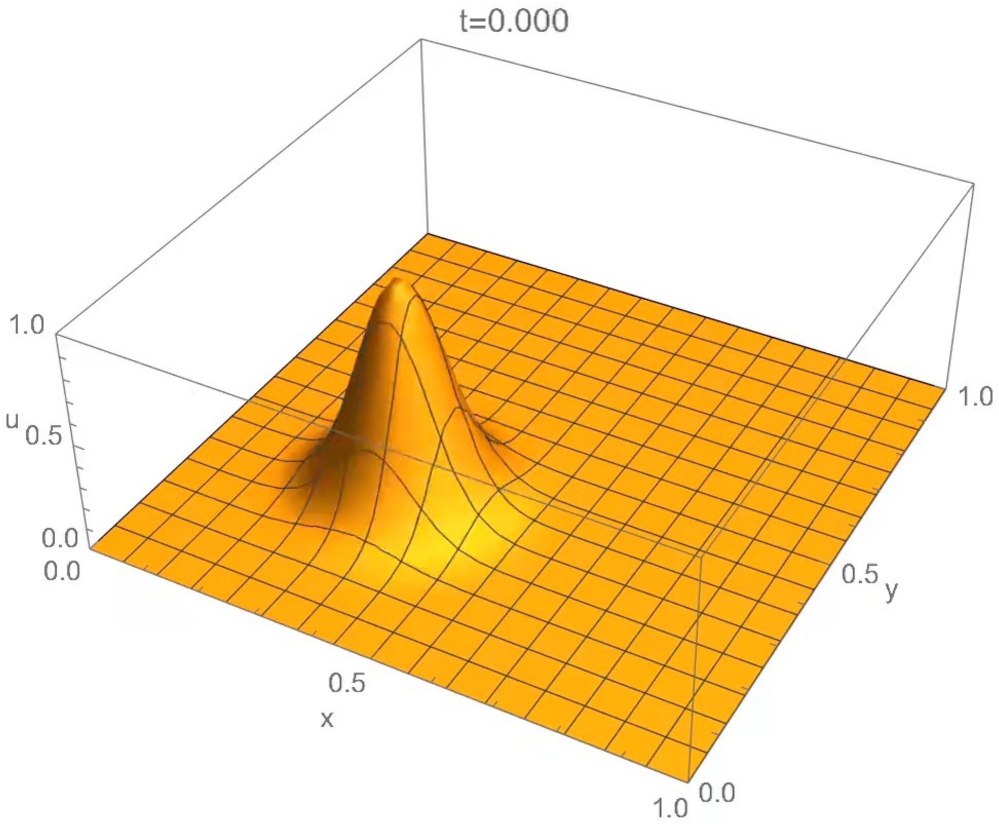
- Consider the Molenkamp-Crowley problem

$$\frac{\partial u}{\partial t} + a \cdot \nabla u = 0, \quad \text{on } \Omega = [0,1]^2,$$
$$u = 0, \quad \text{on } \partial\Omega,$$

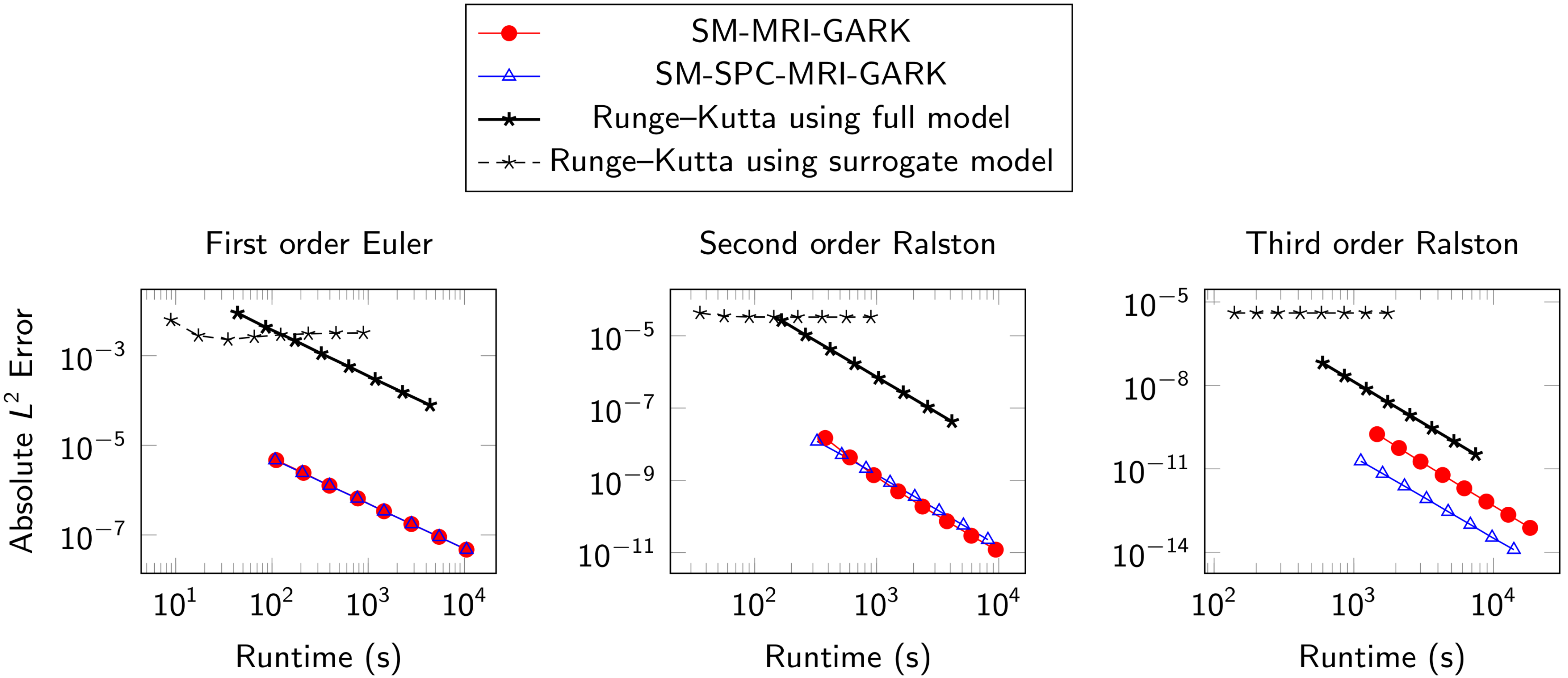
  with the circular wind profile $a(x, y)$.

- f corresponds to a discontinuous Galerkin discretization on a 100×100 uniform triangular mesh, while f_{sur} uses a 50×50 mesh.

- V and W^* are sparse interpolation operators.



Numerical experiment: DG advection



Conclusions

- New methods extend traditional Runge-Kutta and linear multistep methods to incorporate information from a surrogate model.
- This work broadens the scope and applicability of multirate methods.
- The quality of the surrogate model does not affect the order of convergence.
- Experiments show large speedups over traditional integrators, especially when V, W^*, and f_{sur} are inexpensive.
- Future work
  - Additional testing of methods based on linear multistep methods
  - Support for surrogate models that are flow maps instead of ODEs

Questions?

See my website for additional details
https://people.llnl.gov/roberts115

CASC
Center for Applied Scientific Computing

Lawrence Livermore National Laboratory

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.