A New Multirate Time-Stepping Strategy for ODE Systems Equipped with a Surrogate Model

SIAM CSE

Steven Roberts, Andrey A Popov, Arash Sarshar, Adrian Sandu

Computational Science Laboratory
"Compute the Future!"
Department of Computer Science,
Virginia Polytechnic Institute and State University
Blacksburg, VA 24060

March 5, 2020

Compute the Future!

VIRGINIA TECH

Solving large-scale systems of ODEs

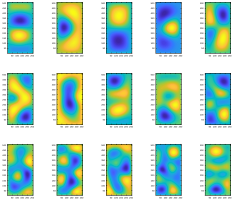
- Consider the system of ordinary differential equations (ODEs)

$$y' = f(y), \qquad y(t_0) = y_0, \qquad y \in \mathbb{C}^N.$$

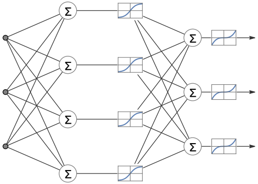
- We will limit our focus to nonstiff systems.
- In many scientific applications, f can be prohibitively expensive and the dimension N intractably large.
- This is common in method of lines discretized PDEs that require a fine spatial mesh.

What about surrogate models?

- In many cases, we can find a cheap but inaccurate surrogate model that approximates the full model f.
- Examples:



(a) Model Order Reduction



(b) Neural Network

(c) Coarser Meshes

- Integrating the surrogate model alone may lead to stability, accuracy, and convergence issues.

The best of both worlds: combining full and surrogate models

- How can we incorporate a surrogate model into the integration of f to reduce expensive full model evaluations?
- In the context of optimization, surrogate models have proved successful in reducing evaluations of expensive objective functions.
- There are some related ideas that can be used in the time-stepping context:
  - Rosenbrock-W methods admit approximate Jacobians
  - Coupling of model order reduction and multirate methods[1]
  - Projective integration
  - High-order/low-order (HOLO) algorithms[2]

[1] Hachtel, Kerler-Back, Bartel, Günther, and Stykel, "Multirate DAE/ODE-Simulation and Model Order Reduction for Coupled Field-Circuit Systems"; Bannenberg, Ciccazzo, and Günther, "Coupling of model order reduction and multirate techniques for coupled dynamical systems".

[2] Chacón et al., "Multiscale high-order/low-order (HOLO) algorithms and applications".

Surrogate model definition

- Suppose we have a surrogate model

$$y'_{\mathrm{sur}} = f_{\mathrm{sur}}(y_{\mathrm{sur}}), \qquad y_{\mathrm{sur}} \in \mathbb{C}^{S}.$$

- Recall that the full model was

$$y' = f(y), \qquad y \in \mathbb{C}^{N}.$$

- The surrogate model could evolve in lower-dimensional subspace: $S < N$.
- To move between the full and surrogate model spaces, we use $\mathrm{V}, \mathrm{W} \in \mathbb{C}^{N \times S}$:

$$y_{\mathrm{sur}} = \mathrm{W}^* y, \qquad y \approx \mathrm{V} y_{\mathrm{sur}}, \qquad \mathrm{W}^* \mathrm{V} = \mathrm{I}_{S \times S}.$$

Using multirating to combine full and surrogate models[3]

- Let's rewrite the original ODE as

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

- 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 model error and is treated with a large timestep.
- The surrogate model guides the solution while the occasional full model evaluation corrects for surrogate model error and ensures convergence.

[3] Roberts, Popov, Sarshar, and Sandu, "A fast time-stepping strategy for ODE systems equipped with a surrogate model".

Multirate methods

- Any multirate method suitable for additively partitioned ODEs could be used.
- Multirate infinitesimal step[4] (MIS) and the multirate infinitesimal GARK[5] (MRI-GARK) extensions make an excellent choice.
- Consider, for example, MIS/MRI Euler applied to $y' = f^{\{\mathfrak{f}\}}(y) + f^{\{\mathfrak{s}\}}(y)$:

$$
\begin{aligned}
v(0) &= y_n, \\
v'(\theta) &= f^{\{\mathfrak{f}\}}(v(\theta)) + f^{\{\mathfrak{s}\}}(y_n), \qquad \text{for } \theta \in [0, H], \\
y_{n+1} &= v(H).
\end{aligned}
$$

- Infinitesimal schemes blur the line between continuous and discrete methods and offer wonderful flexibility in the treatment of $f^{\{\mathfrak{f}\}}$.

[4] Knoth and Wolke, "Implicit-explicit Runge–Kutta methods for computing atmospheric reactive flows"; Wensch, Knoth, and Galant, "Multirate infinitesimal step methods for atmospheric flow simulation".

[5] Sandu, "A Class of Multirate Infinitesimal GARK Methods"; Roberts, Sarshar, and Sandu, "Coupled Multirate Infinitesimal GARK Schemes for Stiff Systems with Multiple Time Scales".

Two types of multirate infinitesimal GARK methods

MRI-GARK:

$$Y_1 = y_n$$

$$\begin{cases} v_i(0) = Y_i, \\ v_i' = \Delta c_i^{\{\mathfrak{s}\}} \, f^{\{\mathfrak{f}\}}(v_i) + \sum_{j=1}^{i+1} \gamma_{i,j}\left(\frac{\theta}{H}\right) f^{\{\mathfrak{s}\}}(Y_j), \\ \qquad \text{for } \theta \in [0, H], \\ Y_{i+1} = v_i(H), \qquad i = 1, \dots, s^{\{\mathfrak{s}\}}, \end{cases}$$

$$y_{n+1} = Y_{s^{\{\mathfrak{s}\}}+1}.$$

Step predictor-corrector MRI-GARK (SPC-MRI-GARK):

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

$$\begin{cases} v(0) = y_n, \\ v' = f^{\{\mathfrak{f}\}}(v) + \sum_{j=1}^{s^{\{\mathfrak{s}\}}} \gamma_j\left(\frac{\theta}{H}\right) f^{\{\mathfrak{s}\}}(Y_j), \\ \qquad \text{for } \theta \in [0, H], \\ y_{n+1} = v(H). \end{cases}$$

Applying MRI-GARK to the ODE

- Now we apply an MRI-GARK method to

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

- This yields the surrogate model MRI-GARK (SM-MRI-GARK) method

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

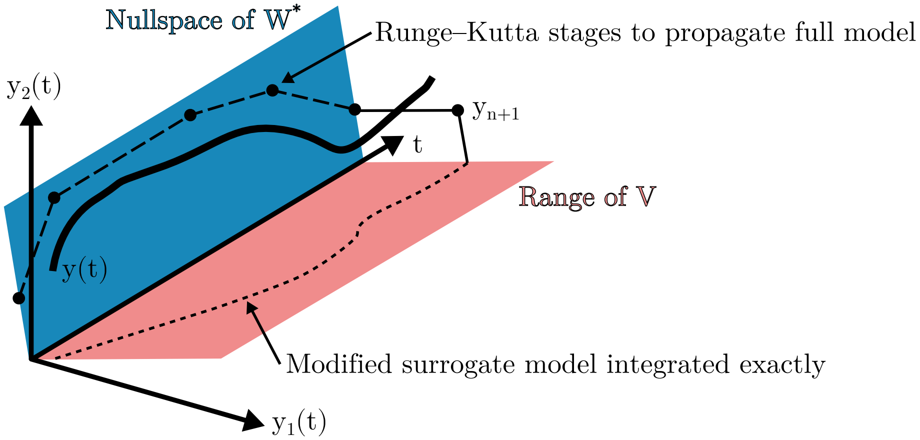
SM-MRI-GARK Euler

- When we use Euler's method as the base method and the coupling $\gamma_{1,1}(t) = 1$, we have the simple method

$$
\begin{aligned}
z(0) &= \mathrm{W}^* y_n \\
z'(\theta) &= f_{\mathrm{sur}}(z(\theta)) + \mathrm{W}^* f(y_n) - f_{\mathrm{sur}}(\mathrm{W}^* y_n), \qquad \text{for } \theta \in [0, H], \\
y_{n+1} &= \mathrm{V} z(H) + (\mathrm{I}_{d \times d} - \mathrm{V}\mathrm{W}^*)(y_n + H f(y_n)).
\end{aligned}
$$

- An ODE is solved in the range of V.
- An Euler step is taken in the nullspace of W^*.

Illustration of SM-MRI-GARK



SM-SPC-MRI-GARK

- We can also apply SPC-MRI-GARK to the ODE to get SM-SPC-MRI-GARK:

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

$$\begin{cases} z(0) = \mathrm{W}^* y_n \in \mathbb{C}^S, \\ z(\theta)' = f_{\mathrm{sur}}(z(\theta)) + \sum_{j=1}^{s^{\{\mathfrak{s}\}}} \gamma_j\left(\frac{\theta}{H}\right)(\mathrm{W}^* f(Y_j) - f_{\mathrm{sur}}(\mathrm{W}^* Y_j)), \\ \qquad \text{for } \theta \in [0, H], \\ y_{n+1} = \mathrm{V} z(H) + (\mathrm{I}_{d \times d} - \mathrm{V}\mathrm{W}^*)\left(y_n + H \sum_{j=1}^{s^{\{\mathfrak{s}\}}} b_j f(Y_j)\right). \end{cases}$$

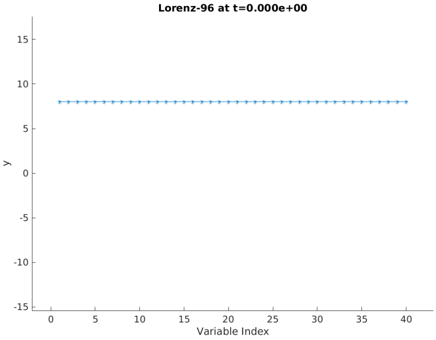
Numerical experiment: Lorenz '96 I

- The Lorenz '96 problem is given by

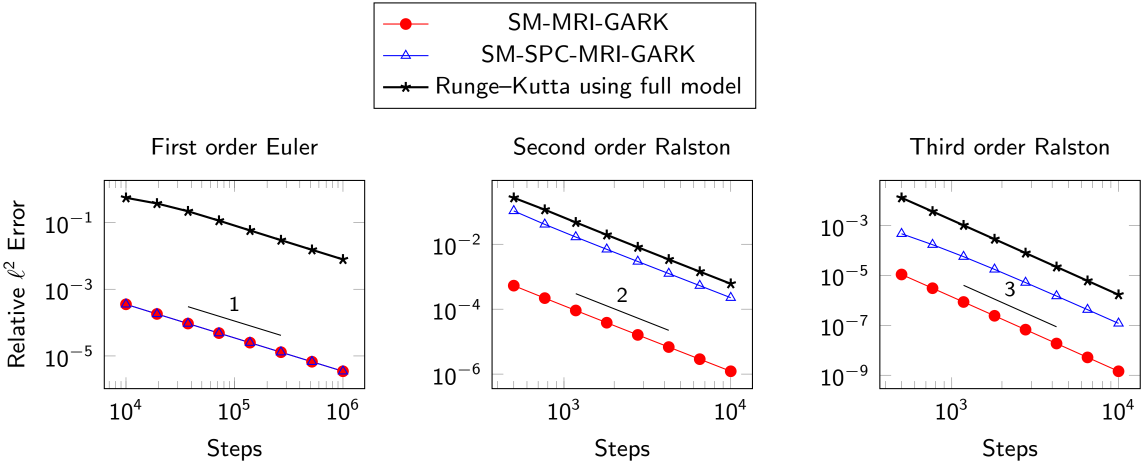
$$\frac{\mathrm{d}}{\mathrm{d}t} X_k = -X_{k-2} X_{k-1} + X_{k-1} X_{k+1} - X_k + F,$$

for $k = 1, 2, \dots, 40$.

- In an offline phase, 5000 snapshots of the state and its derivative were generated over the timespan $[2, 10]$.
- A 3-layer neural network was trained on this data to learn the right-hand side function.
- The neural network is used as f_{sur} and $\mathrm{V} = \mathrm{W} = \mathrm{I}_{40 \times 40}$.



Numerical experiment: Lorenz ’96 II



Numerical experiment: DG advection I

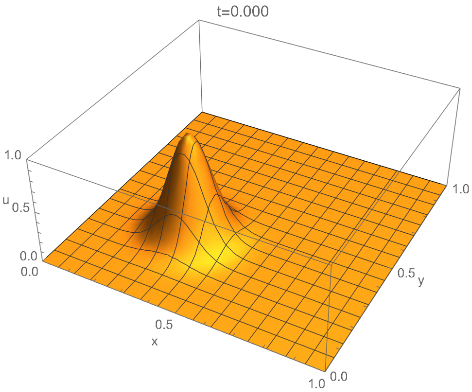
- Consider the Molenkamp–Crowley problem

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

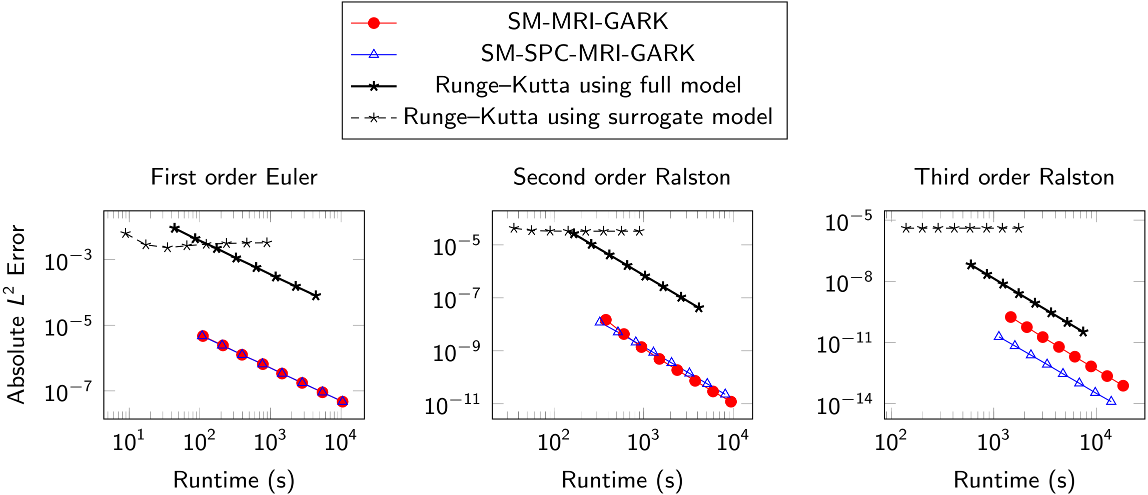
with the circular wind profile

$$a(x,y) = \begin{bmatrix} 2\pi \left(y - \frac{1}{2}\right) \\ -2\pi \left(x - \frac{1}{2}\right) \end{bmatrix}.$$

- 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 II



Conclusions

- Surrogate model MRI-GARK extends traditional Runge–Kutta to incorporate a surrogate model and improve efficiency.
- It combines continuous integration of a surrogate model with discrete evaluations of the full model for error correction.
- Quality of the surrogate model does not affect order of convergence.
- The infinitesimal characteristic allows any method to be applied to the surrogate model.
- Large speedups over traditional Runge–Kutta methods can be achieved, especially when evaluations of V, W*, and f_{sur} are cheap.

References

M.W.F.M. Bannenberg, A. Ciccazzo, and M. Günther. "Coupling of model order reduction and multirate techniques for coupled dynamical systems". In: *Applied Mathematics Letters* 112 (2021), p. 106780. ISSN: 0893-9659. DOI: 10.1016/j.aml.2020.106780.

L. Chacón et al. "Multiscale high-order/low-order (HOLO) algorithms and applications". In: *Journal of Computational Physics* 330 (2017), pp. 21–45. ISSN: 10902716. DOI: 10.1016/j.jcp.2016.10.069.

Christoph Hachtel et al. "Multirate DAE/ODE-Simulation and Model Order Reduction for Coupled Field-Circuit Systems". In: *Scientific Computing in Electrical Engineering*. Ed. by Ulrich Langer, Wolfgang Amrhein, and Walter Zulehner. Cham: Springer International Publishing, 2018, pp. 91–100. ISBN: 978-3-319-75538-0.

O. Knoth and R. Wolke. "Implicit-explicit Runge–Kutta methods for computing atmospheric reactive flows". In: *Applied Numerical Mathematics* 28.327–341 (1998).

Steven Roberts, Arash Sarshar, and Adrian Sandu. "Coupled Multirate Infinitesimal GARK Schemes for Stiff Systems with Multiple Time Scales". In: *SIAM Journal on Scientific Computing* 42.3 (2020), A1609–A1638. DOI: 10.1137/19M1266952.

Steven Roberts et al. "A fast time-stepping strategy for ODE systems equipped with a surrogate model". In: *arXiv preprint arXiv:2011.03688* (2020).

Adrian Sandu. "A Class of Multirate Infinitesimal GARK Methods". In: *SIAM Journal on Numerical Analysis* 57.5 (2019), pp. 2300–2327. DOI: 10.1137/18M1205492.

Jörg Wensch, Oswald Knoth, and Alexander Galant. "Multirate infinitesimal step methods for atmospheric flow simulation". In: *BIT Numerical Mathematics* 49.2 (2009), pp. 449–473. ISSN: 00063835. DOI: 10.1007/s10543-009-0222-3.

Questions?

- Preprint available at:
  https://arxiv.org/abs/2011.03688
- These slides are available on my website:
  https://steven-roberts.github.io